

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 14, 2003, 10:11:42 ; Search time 84 Seconds

(without alignments)  
861.659 Million cell updates/sec

Title: US-09-901-572A-3

Perfect score: 2324

Sequence: 1 MHYFRRCIFPLIVILYGTN.....SSNNNNADKIPGRRPPTFL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2317	99.7	456	AAW36050	Hybrid Marek's dis
2	2004	86.2	1086	AAW36051	Hybrid Marek's dis
3	1917	82.5	615	AAW63230	Mycoplasma gallise
4	1849	79.6	610	AAW63229	Mycoplasma gallise
5	1619	69.7	368	AAW44493	Mycoplasma gallise
6	1619	69.7	368	AAW63227	Mycoplasma gallise
7	1611	63.3	368	AAW76855	Mycoplasma gallise
8	1149	49.4	235	AAW93646	Amino acid sequenc
9	1149	49.4	235	AAW05081	MG-1 antigen. Myc

10	1149	49.4	261	11	AAW05082	TMG-1 antigen. My
11	1149	49.4	261	16	AAW79911	M.gallisepticum 26
12	1124	48.4	261	10	AAW39959	Amino acid (AA) se
13	803	34.6	661	15	AAW63226	Mycoplasma gallise
14	803	34.6	661	16	AAW79910	M.gallisepticum 66
15	711.5	30.6	647	16	AAW11978	Mycobacterium gall
16	699	30.1	648	15	AAW56973	PMGA 1.2 protein o
17	386.5	16.6	183	10	AAW63459	Amino acid sequenc
18	386.5	16.6	183	11	AAW06439	MG-4 antigen. Myc
19	314	13.5	865	14	AAW30169	Marek's Disease V1
20	283	12.2	219	10	AAW93648	Amino acid sequenc
21	283	12.2	219	11	AAW06438	MG-3 antigen. Myc
22	183.5	7.9	1095	22	AAW63030	S. epidermidis ope
23	181.5	7.8	10182	23	ABP38314	Staphylococcus epi
24	177.5	7.6	933	24	ABJ18947	Pathogen specific
25	176.5	7.6	496	23	AAW54490	S. aureus antigen
26	176.5	7.6	933	21	AAW58435	Staphylococcus aur
27	176.5	7.6	933	22	AAW59508	Staphylococcus aur
28	176.5	7.6	936	18	AAW89801	Staphylococcus aur
29	173	7.4	520	23	AAW29262	Staphylococcus aur
30	170	7.3	2086	22	AAW4143	Staphylococcus aur
31	170	7.3	5795	22	AAW37017	Staphylococcus aur
32	169.5	7.3	6281	24	AAW37403	Staphylococcus aur
33	167.5	7.2	10498	22	ABU39119	Pathogen specific
34	166.5	7.2	2659	23	AAW5489	S. aureus antigen
35	162.5	7.0	1237	18	AAW5640	H. pylori ORF 04ep
36	162.5	7.0	1237	20	AAW17187	H. pylori outer me
37	159	6.8	1029	22	AAW4389	Staphylococcus aur
38	159	6.8	1048	22	AAW37490	Staphylococcus aur
39	157.5	6.8	1072	23	ABW54963	Staphylococcus aur
40	156	6.7	807	21	AAW18311	Lactococcus lactis
41	156	6.7	2434	22	AAW4339	Plasmodium falcipa
42	154.5	6.6	2478	22	AAW4320	Staphylococcus aur
43	154.5	6.6	2478	22	AAW37374	Staphylococcus aur
44	154.5	6.6	2478	24	ABJ19002	Staphylococcus aur
45	153.5	6.6	3158	22	AAW37018	Pathogen specific

## ALIGNMENTS

RESULT 1  
AAW36050 standard; Protein; 456 AA.

XX AC AAW36050;

XX DT 15-JUL-1998 (first entry)

XX DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein; antigen; vaccine; poultry.

XX OS Chimeric - Marek's disease gammaherpesvirus.

XX OS Chimeric - Mycoplasma gallisepticum.

XX FT Key

XX FT Region

XX FT Region

XX FT Region

XX FT Region

XX FT Region

XX FT Region

XX FT Region

XX FT Region

XX FT Region

PI Salto S, Tsuzaki Y, Yanagida N;  
 XX WPI; 1997-503046/46.  
 DR N-PSDB; AAT96595.  
 XX  
 PT Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 16-19; 51pp; Japanese.  
 XX  
 CC This sequence represents the chimeric protein 40 K-S which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 CC  
 XX  
 SQ Sequence 456 AA;  
 Query Match 99.7%; Score 2317; DB 18; Length 456;  
 Best Local Similarity 99.6%; Pred. No. 5.2e-143;  
 Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MHYFRNCIFELIVILYIGTNSPSTQNTSRVSVQSEESTFYLCPPVGSTVRL 60  
 DB 1 MHYFRNCIFELIVILYIGTNSPSTQNTSRVSVQSEESTFYLCPPVGSTVRL 60  
 QY 61 EFGCMSITKCDANPNNGQTOLEARMELTDLINAKMTLASLDYAKIESLSAYSEAE 120  
 DB 61 EFGCMSITKCDANPNNGQTOLEARMELTDLINAKMTLASLDYAKIESLSAYSEAE 120  
 QY 121 TVNNNINATLEOLKAKTNLESAINQANTDKTFDNEHPNVEAYKALKTTLEQRATNLE 180  
 DB 121 TVNNNINATLEOLKAKTNLESAINQANTDKTFDNEHPNVEAYKALKTTLEQRATNLE 180  
 QY 181 GLSSTAYNOIRNNLVLDYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEQ 240  
 DB 181 GLSSTAYNOIRNNLVLDYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEQ 240  
 QY 241 KTNADALNSNFIKVIQONNEQSFVGTFTNANVOPSNYSFVAFSADVPVNYKYARRTYWN 300  
 DB 241 KTNADALNSNFIKVIQONNEQSFVGTFTNANVOPSNYSFVAFSADVPVNYKYARRTYWN 300  
 QY 301 GDEPSRIILANTNSITDVSMTIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360  
 DB 301 GDEPSRIILANTNSITDVSMTIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360  
 QY 361 GLOYKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 DB 361 GLOYKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 QY 421 EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 DB 421 EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 RESULT 2  
 ID AAM36051 standard; Protein; 1086 AA.  
 XX AAM36051;  
 AC  
 XX  
 DT 15-UTR-1998 (first entry)  
 XX  
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.  
 XX  
 KM Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
 XX antigen; vaccine; poultry.  
 KM  
 OS Chimeric - Marek's disease gammaherpesvirus.  
 OS Chimeric - Mycoplasma gallisepticum.  
 XX

FH Key Location/Qualifiers  
 FT Region 1..672 /note="derived from Marek's disease virus gB protein"  
 FT Region 693..1086 /note="derived from M. gallisepticum antigen"  
 FT Region /note="derived from M. gallisepticum antigen"  
 XX  
 XX W09736924-A1.  
 XX  
 XX 09-OCT-1997.  
 XX  
 XX 28-MAR-1997; 97WO-JP01084.  
 XX  
 XX 29-MAR-1996; 96JP-0103548.  
 XX  
 XX (UAPG) NIPPON ZEON KK.  
 XX  
 PI Salto S, Tsuzaki Y, Yanagida N;  
 XX WPI; 1997-503046/46.  
 DR N-PSDB; AAT96596.  
 XX  
 PT Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 22-30; 51pp; Japanese.  
 XX  
 CC This sequence represents the chimeric protein 40 K-C which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 CC  
 XX  
 SQ Sequence 1086 AA;  
 Query Match 86.2%; Score 2004; DB 18; Length 1086;  
 Best Local Similarity 99.2%; Pred. No. 4.2e-122;  
 Matches 393; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 61 EFGCMSITKCDANPNNGQTOLEARMELTDLINAKMTLASLDYAKIESLSAYSEAE 120  
 DB 61 EFGCMSITKCDANPNNGQTOLEARMELTDLINAKMTLASLDYAKIESLSAYSEAE 120  
 QY 121 TVNNNINATLEOLKAKTNLESAINQANTDKTFDNEHPNVEAYKALKTTLEQRATNLE 180  
 DB 121 TVNNNINATLEOLKAKTNLESAINQANTDKTFDNEHPNVEAYKALKTTLEQRATNLE 180  
 QY 181 GLSSTAYNOIRNNLVLDYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEQ 240  
 DB 181 GLSSTAYNOIRNNLVLDYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEQ 240  
 QY 241 KTNADALNSNFIKVIQONNEQSFVGTFTNANVOPSNYSFVAFSADVPVNYKYARRTYWN 300  
 DB 241 KTNADALNSNFIKVIQONNEQSFVGTFTNANVOPSNYSFVAFSADVPVNYKYARRTYWN 300  
 QY 301 GDEPSRIILANTNSITDVSMTIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360  
 DB 301 GDEPSRIILANTNSITDVSMTIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360  
 QY 361 GLOYKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 DB 361 GLOYKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 QY 421 EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 DB 1051 EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGFTL 1086  
 RESULT 3  
 ID AAR63230 standard; Protein; 615 AA.  
 XX AAR63230  
 XX

```

AC AAR63230;
XX 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
XX Mycoplasma gallisepticum antigen (UM-67).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
XX OS
XX Key Location/Qualifiers
FH Protein 1..615
FT /note= "Trp residues correspond to TGA codons"
XX
XX MO9423019-A1.
XX 13-OCT-1994.
XX 31-MAR-1994; 94WO-JP00541.
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saito, Takahashi K;
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77857.
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 87-91; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX ORF encodes an antigenic polypeptide (AAR63230). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 615 AA;
SQ
Query Match 82.5%; Score 1917; DB 15; Length 615;
Best Local Similarity 98.7%; Pred. No. 9.2e-117;
Matches 378; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 64 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 123
DB 27 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 86
QY 124 NNUNATLEOLKMAKTNLESAINQANTDCTFDNEHPNVVEAYKAKTTLLEORATNLGSL 183
DB 87 NNUNATLEOLKMAKTNLESAINQANTDCTFDNEHPNVVEAYKAKTTLLEORATNLGSL 146
QY 184 STAYQIRNNLVLDLYNKASSLITKTLDPNGTLLDSNEITTVNINNTLSTINEOKTN 243
DB 147 STAYQIRNNLVLDLYNKASSLITKTLDPNGTLLDSNEITTVNINNTLSTINEOKTN 206
QY 244 ADAISNSIKKYIONNEOSFVGTFNNANVOPSNYSFVAPADVTPEVNYKYARIVWNGDE 303
DB 207 ADAISNSIKKYIONNEOSFVGTFNNANVOPSNYSFVAPADVTPEVNYKYARIVWNGDE 266
QY 304 PSSRLIANTNSTIDVSWIYSLAGTNTKYOFSSNSYSPGTGYLYFPYKYVKADANNVGLQ 363
DB 267 PSSRLIANTNSTIDVSWIYSLAGTNTKYOFSSNSYSPGTGYLYFPYKYVKADANNVGLQ 326
QY 364 YKLNNGNVOVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGECN 423

```

```

DB 327 YKLNNGNVOVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGECN 386
QY 424 MNKYAPMIGNTYILSSNNENADKI 446
DB 387 MNKYAPMIGNTYILSSNNENADKI 409
RESULT 4
ID AAR63229 standard; Protein; 610 AA.
XX AAR63229;
XX 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
XX Mycoplasma gallisepticum antigen (UM-66).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
XX OS
XX Key Location/Qualifiers
FH Protein 1..610
FT /note= "Trp residues correspond to TGA codons"
XX
XX MO9423019-A1.
XX 13-OCT-1994.
XX 31-MAR-1994; 94WO-JP00541.
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saito, Takahashi K;
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77856.
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 610 AA;
SQ
Query Match 79.6%; Score 1849; DB 15; Length 610;
Best Local Similarity 95.0%; Pred. No. 2.5e-112;
Matches 363; Conservative 11; Mismatches 8; Indels 0; Gaps 0;
QY 64 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 123
DB 27 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 86
QY 124 NNUNATLEOLKMAKTNLESAINQANTDCTFDNEHPNVVEAYKAKTTLLEORATNLGSL 183
DB 87 NNUNATLEOLKMAKTNLESAINQANTDCTFDNEHPNVVEAYKAKTTLLEORATNLGSL 146
QY 184 STAYQIRNNLVLDLYNKASSLITKTLDPNGTLLDSNEITTVNINNTLSTINEOKTN 243

```

```

Db 147 STAYNOIRNNVLDLYNKASLITKTLDPINGGTLIDSLNEITTVNINNTLSTINEOKTN 206
QY 244 ADALNSFTKVIQONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 303
Db 207 ADALNSFTKVIQONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 266
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNVGLQ 363
Db 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNVGLQ 326
QY 364 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEN 423
Db 327 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEN 386
QY 424 NKKVAPMIGNIYISSNENNAADK 445
Db 387 NKKVAPMIGNIYISSNENNAADK 408

```

## RESULT 5

AA044493 ID AA044493 standard; Protein; 368 AA.

AA044493; AC XX

25-MAR-2003 (updated)  
16-JUN-1994 (first entry)

DE Mycoplasma gallisepticum 40kd antigen.

KW Vaccine; mycoplasma infection; poultry; fowl.

OS Mycoplasma gallisepticum.

FN Key Location/Qualifiers

FT Misc-difference 262 /note= "corresponds to NNN codon in AA053419"

FT Misc-difference 283 /note= "corresponds to NNN codon in AA053419"

FN WO9324646-A1.

XX 09-DEC-1993.

PF 28-MAY-1993; 93WO-JP00715.

PR 29-MAY-1992; 92JP-0138819.

PA (JAPG ) NIPPON ZEON KK.

PI (SHIO ) SHIONOGI & CO LTD.

PI Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;

DR WPI; 1993-405837/50.

DR N-PSDB; AA053419.

PT Mycoplasma gallisepticum antigen and DNA coding for it - useful

PT for vaccination of fowl against mycoplasma infections

PS Claim 2; Page 23-26; 37pp; Japanese.

CC The sequence coding for the 40kDa antigen was obtained by PCR

CC amplification of M.gallisepticum genomic DNA. The antigen reacts

CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used

CC as a vaccine to protect fowl from M.gallisepticum infection.

CC (Updated on 25-MAR-2003 to correct FN field.)

XX Sequence 368 AA;

```

QY 64 CMSITKKDANPNNGOTOLEAAMELTDLINAKAMTLASLOPYAKIEASLSAYSFAETVN 123
Db 27 CMSITKKDANPNNGOTOLEAAMELTDLINAKAMTLASLOPYAKIEASLSAYSFAETVN 86
QY 124 NNLANTEBOLKMAKTNLESAINQANTDKTTFDNEHPNIVEYKALKTTLEORATVLEGLS 183
Db 87 NNLANTEBOLKMAKTNLESAINQANTDKTTFDNEHPNIVEYKALKTTLEORATVLEGLS 146
QY 184 STAYNOIRNNVLDLYNKASLITKTLDPINGGTLIDSLNEITTVNINNTLSTINEOKTN 243
Db 147 STAYNOIRNNVLDLYNKASLITKTLDPINGGTLIDSLNEITTVNINNTLSTINEOKTN 206
QY 244 ADALNSFTKVIQONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 303
Db 207 ADALNSFTKVIQONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 266
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNVGLQ 363
Db 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNVGLQ 326
QY 364 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 400
Db 327 YKLNNGNVQVEFATSTSANNTTANPTQOLMRLTLK 363

```

## RESULT 6

AA063227 ID AA063227 standard; Protein; 368 AA.

AA063227; AC XX

25-MAR-2003 (updated)  
23-JUN-1995 (first entry)

DE Mycoplasma gallisepticum 40kd antigen.

KW recombinant avipox virus; live vaccine; mycoplasma 40kd antigen;

OS TTM-1.

OS Mycoplasma gallisepticum.

FN Key Location/Qualifiers

FT Misc-difference 262 /note= "corresponds to a NNN codon"

FT Misc-difference 283 /note= "corresponds to a NNN codon"

FN WO9423019-A1.

PF 31-MAR-1994; 94WO-JP00541.

PR 31-MAR-1993; 93JP-0074139.

PR 30-SEP-1993; 93JP-0245625.

XX (JAPG ) NIPPON ZEON KK.

PI (SHIO ) SHIONOGI & CO LTD.

PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;

DR WPI; 1994-333181/41.

DR N-PSDB; AA077854.

PT Recombinant avipox virus combining DNA encoding a polypeptide -

PT exhibiting antigenicity of mycoplasma, useful for the production

XX of a live vaccine

XX Claim 4; Page 71-74; 123pp; Japanese.

XX The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding



CC for the 40kD antigen of Mycoplasma gallisepticum under the control  
 CC of a synthetic promoter. A 1300 bp restriction fragment containing  
 CC the promoter-ORF sequence was excised and was used in the  
 CC construction of plasmid pN27929-R2. This in turn was involved in the  
 CC construction of a recombinant avipox virus vector comprising the  
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from  
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FPV  
 CC sequences. The recombinant avipox virus is useful as a live vaccine  
 CC to protect against infection by Mycoplasma gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 368 AA;

Query Match 69.7%; Score 1619; DB 15; Length 368;

Best Local Similarity 95.5%; Pred. No. 1.2e-97;  
 Matches 322; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 64 CMSITKKDANPNNGCOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 DB 27 CMSITKKDANPNNGCOTOLEAARMELTDLINAKARLASLODYAKIEASLSAYSEAEVTN 86  
 QY 124 NNINATLEOLKMAKTNLESAINOANTDKTPEDEHPNLVEAYKAKTLEORATNLEGLS 183  
 DB 87 NNINATLEOLKMAKTNLESAINOANTDKTPEDEHPNLVEAYKAKTLEORATNLEGLA 146  
 QY 184 STAYNOIRNNIVDLVYNKASSLITKTLDPLNGGTLDSNEITTVARNINNTLSTINEOKTN 243  
 DB 147 STAYNOIRNNIVDLVYNKASSLITKTLDPLNGGTLDSNEITTVARNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNESQFVGFTFNANVOPSNSYFVAFSADVTPVNYKYARRTVWNGDE 303  
 DB 207 ADALSNFIRKVIQNNESQFVGFTFNANVOPSNSYFVAFSADVTPVNYKYARRTVXNGDE 266  
 QY 304 PSSRIILANTNSITVSWIYSLAGTNTKYOFSESNYGPSTGYLYPPYKLVKADANNVGLQ 363  
 DB 267 PSSRIILANTNSITVSWIYSLAGTNTKYOFSESNYGPSTGYLYPPYKLVKADANNVGLQ 326  
 QY 364 YKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKYAK 400  
 DB 327 YKLNNGNVQOVEFATSTSANNTTANPTQOLMRKLK 363

RESULT 7

AAR76955  
 ID AAR76955 standard; Protein; 368 AA.

XX AAR76955;

XX 25-MAR-2003 (updated)  
 DT 12-MAR-1996 (first entry)

XX Mycoplasma gallisepticum antigenic protein TTM-1.

XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Misc-difference 262 /note= "any amino acid"

FT Misc-difference 283 /note= "any amino acid"

XX JF07133295-A.

XX 23-MAY-1995.

XX 27-AUG-1993; 93UP-0213102.

XX 27-AUG-1993; 93UP-0213102.

XX (UAPG ) NIPPON ZEON KK.

XX (SHIO ) SHIONOGI & CO LTD.

XX WPI; 1995-220782/29.  
 DR N-PSDB; AAQ94711.

XX A new antigenic protein which reacts with Mycoplasma gallisepticum -  
 PT is useful in a component vaccine for use against poultry infected  
 XX with M. gallisepticum.

XX Claim 6; Figs 5-6; 33pp; Japanese.

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein  
 CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum  
 CC infectious diseases in poultry, and as a diagnostic agent for  
 CC M. gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 368 AA;

Query Match 69.3%; Score 1611; DB 16; Length 368;

Best Local Similarity 95.0%; Pred. No. 4.1e-97;  
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 64 CMSITKKDANPNNGCOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 DB 27 CMSITKKDANPNNGCOTOLEAARMELTDLINAKARLASLODYAKIEASLSAYSEAEVTN 86  
 QY 124 NNINATLEOLKMAKTNLESAINOANTDKTPEDEHPNLVEAYKAKTLEORATNLEGLS 183  
 DB 87 NNINATLEOLKMAKTNLESAINOANTDKTPEDEHPNLVEAYKAKTLEORATNLEGLA 146  
 QY 184 STAYNOIRNNIVDLVYNKASSLITKTLDPLNGGTLDSNEITTVARNINNTLSTINEOKTN 243  
 DB 147 STAYNOIRNNIVDLVYNKASSLITKTLDPLNGGTLDSNEITTVARNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNESQFVGFTFNANVOPSNSYFVAFSADVTPVNYKYARRTVWNGDE 303  
 DB 207 ADALSNFIRKVIQNNESQFVGFTFNANVOPSNSYFVAFSADVTPVNYKYARRTVXNGDE 266  
 QY 304 PSSRIILANTNSITVSWIYSLAGTNTKYOFSESNYGPSTGYLYPPYKLVKADANNVGLQ 363  
 DB 267 PSSRIILANTNSITVSWIYSLAGTNTKYOFSESNYGPSTGYLYPPYKLVKADANNVGLQ 326  
 QY 364 YKLNNGNVQOVEFATSTSANNTTANPTPAVDEIKYAK 400  
 DB 327 YKLNNGNVQOVEFATSTSANNTTANPTQOLMRKLK 363

RESULT 8

AAP93646  
 ID AAP93646 standard; Protein; 235 AA.

XX AAP93646;

XX 25-MAR-2003 (updated)  
 DT 11-MAY-1990 (first entry)

XX Amino acid sequence of Mycoplasma gallisepticum (MG1) polypeptide.

XX Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0305441.

XX 31-MAY-1989; 89EP-0305441.

XX 02-JUN-1988; 88UP-0136343.

XX (UAPG ) NIPPON ZEON KK.

XX (SHIO ) SHIONOGI SEIYAKU KK.

PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI SEIYAKU KK.  
 XX  
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 XX  
 DR WPI; 1989-358393/49.  
 DR N-PSDB; AAN92568.  
 XX  
 PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 PT vaccines.  
 XX  
 PS Disclosure; Fig.1a; 31pp; English.  
 CC This amino acid sequence of MG1 is encoded by M1 DNA and elicits an  
 CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a  
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro  
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 235 AA;  
 Query Match 49.4%; Score 1149; DB 10; Length 235;  
 Best Local Similarity 97.9%; Pred. No. 2.9e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 64 CMSITKKDANPNNGOTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 123  
 DB 1 CMSITKKDANPNNGOTOLEAARMELTDLINAKARTLASLDYAKIEASLSAYSEAEVTN 60  
 QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 61 NNLAATLEQLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 120  
 QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 121 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 180  
 QY 244 ADALNSPIKVIQNNESFVGTFTNNAVOPSNYSFVAFSADVTVPVNYKARRTV 298  
 DB 181 ADALNSPIKVIQNNESFVGTFTNNAVOPSNYSFVAFSADVTVPVNYKARRTV 235  
 RESULT 9  
 AAR05081  
 ID AAR05081 standard; protein; 235 AA.  
 XX  
 AC AAR05081;  
 XX  
 DT 10-MAR-2003 (updated)  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE MG-1 antigen.  
 XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.  
 XX  
 OS Mycoplasma gallisepticum.  
 OS JP02111795-A.  
 XX  
 PD 24-APR-1990.  
 XX  
 PF 02-JUN-1989; 89JP-0136343.  
 XX  
 PR 02-JUN-1989; 89JP-0136343.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI KK.  
 XX  
 DR WPI; 1990-169109/22.  
 DR N-PSDB; AAQ04686.

XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes  
 PT antigen protein of the disease and recombinant vector  
 PT incorporated with its coding gene.  
 XX  
 PS Claim 2; Fig 1a; 20pp; Japanese.  
 CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.  
 CC See also AAR05081-2 and AAR06437-41.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)  
 XX  
 SQ Sequence 235 AA;  
 Query Match 49.4%; Score 1149; DB 11; Length 235;  
 Best Local Similarity 97.9%; Pred. No. 2.9e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 64 CMSITKKDANPNNGOTOLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVTN 123  
 DB 1 CMSITKKDANPNNGOTOLEAARMELTDLINAKARTLASLDYAKIEASLSAYSEAEVTN 60  
 QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 61 NNLAATLEQLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 120  
 QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 121 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 180  
 QY 244 ADALNSPIKVIQNNESFVGTFTNNAVOPSNYSFVAFSADVTVPVNYKARRTV 298  
 DB 181 ADALNSPIKVIQNNESFVGTFTNNAVOPSNYSFVAFSADVTVPVNYKARRTV 235  
 RESULT 10  
 AAR05082  
 ID AAR05082 standard; protein; 261 AA.  
 XX  
 AC AAR05082;  
 XX  
 DT 10-MAR-2003 (updated)  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE TMG-1 antigen.  
 XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.  
 XX  
 OS Mycoplasma gallisepticum.  
 OS JP02111795-A.  
 XX  
 PD 24-APR-1990.  
 XX  
 PF 02-JUN-1989; 89JP-0136343.  
 XX  
 PR 02-JUN-1989; 89JP-0136343.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI KK.  
 XX  
 DR WPI; 1990-169109/22.  
 DR N-PSDB; AAQ04687.  
 XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes  
 PT antigen protein of the disease and recombinant vector  
 PT incorporated with its coding gene.  
 XX  
 PS Claim 2; Fig 2; 20pp; Japanese.

CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.  
 CC See also AAR05081 and AAR06437-41.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)

XX Sequence 261 AA;

Query Match 49.4%; Score 1149; DB 11; Length 261;  
 Best Local Similarity 97.9%; Pred. No. 3.4e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGOTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEIVN 123  
 DB 27 CMSITKQDANPNNGOTQLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEIVN 86  
 QY 124 NNLAATLEQLMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEORATNLEGLS 183  
 DB 87 NNLAATLEQLMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEORATNLEGLA 146  
 QY 184 STAYNQIRNNLVLDLYNKAASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 147 STAYNQIRNNLVLDLYNKAASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 206  
 QY 244 ADALNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 298  
 DB 207 ADALNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 261

#### RESULT 11

AAR79911 standard; Protein; 261 AA.

XX ID AAR79911 standard; Protein; 261 AA.  
 XX AC AAR79911;  
 XX DT 19-JUL-1996 (first entry)  
 XX DE M.gallisepticum 261 amino acid protein.  
 XX KM Detection; probe; primer; PCR; amplification; secretion; lung;  
 XX KW avian chronic respiratory disease; respiratory tract; nasal cavity.  
 XX OS Mycoplasma gallisepticum.  
 XX JN JP07236498-A.  
 XX PD 12-SEP-1995.  
 XX PF 25-FEB-1994; 94JP-0052764.  
 XX PR 25-FEB-1994; 94JP-0052764.  
 XX PA (JAPG ) NIPPON ZEON KK.  
 XX PA (SHIO ) SHIONOGI & CO LTD.  
 XX WP1; 1995-347462/45.  
 XX DR N-PSDB; AAT04076.  
 XX PT Detection of Mycoplasma gallisepticum - for the quick detection,  
 XX PT i.e. within one day, of avian chronic respiratory diseases  
 XX PS Claim 3; Page 10-11; 11pp; Japanese.  
 XX CC This is the amino acid sequence of a 261 amino acid protein encoded  
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid  
 CC protein) can be used to detect M.gallisepticum using probes based on  
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 CC complement of bases 893-919, 1908-1934 and the complement of bases  
 CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA  
 CC encoding this protein. The method using these sequences is faster i.e. is

CC able to detect M.gallisepticum, which causes avian chronic respiratory  
 CC diseases, within one day, from avian secretions, washings from the lung,  
 CC respiratory tract, nasal cavity, etc.

XX Sequence 261 AA;

Query Match 49.4%; Score 1149; DB 16; Length 261;  
 Best Local Similarity 97.9%; Pred. No. 3.4e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGOTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEIVN 123  
 DB 27 CMSITKQDANPNNGOTQLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEIVN 86  
 QY 124 NNLAATLEQLMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEORATNLEGLS 183  
 DB 87 NNLAATLEQLMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKTTLEORATNLEGLA 146  
 QY 184 STAYNQIRNNLVLDLYNKAASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 147 STAYNQIRNNLVLDLYNKAASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 206  
 QY 244 ADALNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 298  
 DB 207 ADALNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 261

#### RESULT 12

AAP93959 standard; Protein; 261 AA.

XX ID AAP93959 standard; Protein; 261 AA.  
 XX AC AAP93959;  
 XX DT 25-MAR-2003 (updated)  
 XX DT 11-MAY-1990 (first entry)  
 XX DE Amino acid (AA) sequence of TWG-1 polypeptide.  
 XX KM Mycoplasma gallisepticum; Poultry vaccine; ss;  
 XX OS Mycoplasma gallisepticum.  
 XX JN EP345021-A.  
 XX PD 06-DEC-1989.  
 XX PF 31-MAY-1989; 89EP-0305441.  
 XX PF 31-MAY-1989; 89EP-0305441.  
 XX PR 02-JUN-1988; 88JP-0136343.  
 XX PA (JAPG ) NIPPON ZEON KK.  
 XX PA (SHIO ) SHIONOGI SEIYAKU KK.  
 XX PA (JAPG ) NIPPON ZEON KK.  
 XX PA (SHIO ) SHIONOGI SEIYAKU KK.  
 XX PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 XX PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 XX DR WP1; 1989-358393/49.  
 XX DR N-PSDB; AAN92574.  
 XX PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 XX PT vaccines.  
 XX PS Disclosure; Fig.2; 31pp; English.  
 XX CC This AA sequence of TWG-1 is encoded by TW-1 base sequence. It has the  
 CC same sequence as that of a polypeptide expressed in Mycoplasma  
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted  
 CC into a recombinant vector used to transform a host the antigen protein  
 CC produced can be used as a vaccine to prevent and diagnose MG infection.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 261 AA;  
 Query Match 48.4%; Score 1124; DB 10; Length 261;  
 Best Local Similarity 97.0%; Pred. No. 1.4e-65;  
 Matches 229; Conservative 2; Mismatches 3; Indels 2; Gaps 2;  
 QY 64 CMSITRKQANPNNGQTOLEAAMELTDLINAKAMTLASIQDYAKIEASISSAYSSEATYN 123  
 DB 27 CMSITRKQANPNNGQTOLEAAMELTDLINAKARLASIQDYAKIEASISSAYSSEATYN 86  
 QY 124 NNUNATLEOLKMAKTNLSAIAQANTDKTTPNEHPNLVEAYKALKTTLEQATNLEGLS 183  
 DB 87 NNUNATLEOLKMAKTNLSAIAQANTDKTTPNEHPNLVEAYKALKTTLEQATNLEGLA 146  
 QY 184 STAYNOIRNNLVLDYNNKASLITTKTLDPLNGGTLDSNEITTVNRNINNTLSTINE-OKT 242  
 DB 147 STAYNOIRNNLVLDYNNKASLITTKTLDPLNGGTLDSNEITTVNRNINNTLSTINEOQKT 205  
 QY 243 NADALNSFIKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTPVNYKARRTV 298  
 DB 206 NADALNSFIKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTPVNYKARRTV 261  
 RESULT 13  
 AAR63226  
 ID AAR63226 standard; Protein; 661 AA.  
 XX  
 AC AAR63226;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 23-JUN-1995 (first entry)  
 XX  
 DE Mycoplasma gallisepticum antigen (UM-81).  
 XX  
 KM recombinant avipox virus; live vaccine; mycoplasma antigen.  
 XX  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..661  
 FT /note= "Trp residues correspond to TGA codons"  
 XX  
 PN W09423019-A1.  
 XX  
 PD 13-OCT-1994.  
 XX  
 PF 31-MAR-1994; 94MO-JP00541.  
 XX  
 PR 31-MAR-1993; 93JP-0074139.  
 PR 30-SEP-1993; 93JP-0245625.  
 XX  
 PA (JAPG) NIPPON ZEON KK.  
 PA (SHIO) SHIONOGI & CO LTD.  
 XX  
 PI Aoyama S, Funato H, Tritani Y, Ohkawa S, Ohsawa T;  
 PI Saeki S, Saitos, Takahashi K;  
 XX  
 DR WPI; 1994-333181/41.  
 DR N-PSDB; AAQ77853.  
 XX  
 PT Recombinant avipox virus combining DNA encoding a polypeptide -  
 PT exhibiting antigenicity of mycoplasma, useful for the production  
 PT of a live vaccine  
 XX  
 PS Claim 4; Page 61-65; 123pp; Japanese.  
 XX  
 CC A restriction fragment of the insert of M.gallisepticum genomic clone  
 CC PUM-81 containing an open reading frame was sequenced (AAQ77853). The  
 CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox  
 CC virus comprising the coding sequence can be used as a live vaccine to

CC Protect against infection by Mycoplasma gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 661 AA;  
 Query Match 34.6%; Score 803; DB 15; Length 661;  
 Best Local Similarity 41.4%; Pred. No. 4.2e-44;  
 Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;  
 QY 51 PPVGSIVIRLEFGCMSTIKQANPNNGQTOLEAAMELT-----LINAKMTLASIQ 103  
 DB 35 PPSGG-----NNNGGNTNPSDGGMNAAKELADAKAALTTLINETSANLASYE 84  
 QY 104 DYAKIEASISSAYSSEAEVNNNNLVEOLKMAKTNLSAIAQANTDKTTPNEHPNLVE 163  
 DB 85 DYAKIEASISSAYSSEAEVNNNNLVEOLKMAKTNLSAIAQANTDKTTPNEHPNLVE 144  
 QY 164 AYKALKTTLEQATNLEGLSSTAYNOIRNNLVLDYNNKASLITTKTLDPLNGGTLDSNEI 223  
 DB 145 AYNNLKETLKEKTNLDSLAEENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPVMSV 203  
 QY 224 TTVNRNINNTLSTINEQKTNDALNSFIKVIQNNESFVGTFTNANVOPSNYSFVAFS 283  
 DB 204 TQANQDITNATSRLIAMKQNDNLANSFVKSLVKNMLRVDVANNOE-QPANTSFVGS 262  
 QY 284 ADVTPVNYKARRTVYNGDE-----PSSRIIANTNSITDVSWIYSLAGTNTKQFSPSN 337  
 DB 263 VNVDTPNMNFQKRYKVASSENTPLATTPAEDTQQAASLTIDVSIYSLNGAKAKTISRY 322  
 QY 338 YG-PSTGYLYFPYKLVKADANNVGLQYKLNNGVQVEF-----ATSTSAN--NTTA 387  
 DB 323 FGAERTALVLPYKLVKXSD--NVGLQYKLNNGGDTKQINFVQTPASGSSDVAANEETMA 380  
 QY 388 NP-----TPAVDEIKVAKIVYSGRFGQNTIEISVPTGEGMNNKVAPEIGNIYLS---- 438  
 DB 381 SPAEQSAPVTDIDIKIAVVALSNLKFNSNTIEFSVPTG-----KAAPMIGMNYILSNSE 435  
 QY 439 -----NENN-----ADKIPGV 449  
 DB 436 VNRKKTIDDLFGNSFNENNPTAVTVDLKGY 467  
 RESULT 14  
 AAR79910  
 ID AAR79910 standard; Protein; 661 AA.  
 XX  
 AC AAR79910;  
 XX  
 DT 19-JUL-1996 (first entry)  
 DT XX  
 DE M.gallisepticum 661 amino acid protein.  
 XX  
 KM Detection; probe; primer; PCR; amplification; secretion; lung;  
 KM avian chronic respiratory disease; respiratory tract; nasal cavity.  
 XX  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 220  
 FT /note= "encoded by TGA"  
 FT Misc-difference 270  
 FT /note= "encoded by TGA"  
 FT Misc-difference 305  
 FT /note= "encoded by TGA"  
 FT FT Misc-difference 581  
 FT /note= "encoded by TGA"  
 XX  
 PN JP07236498-A.  
 XX  
 PD 12-SEP-1995.  
 XX  
 PR 25-FEB-1994; 94JP-0052764.

PR 25-FEB-1994; 94JP-0052764.  
 XX (JAP) NIPPON ZEON KK.  
 PA (SHIO) SHIONOGI & CO LTD.  
 XX  
 DR WPI; 1995-347462/45.  
 DR N-PSDB; AAT04075.  
 XX  
 PT Detection of Mycoplasma gallisepticum - for the quick detection,  
 PT i.e. within one day, of avian chronic respiratory diseases  
 XX  
 PS Claim 2; Page 7-10; 11pp; Japanese.  
 XX  
 CC This is the amino acid sequence of a 661 amino acid protein encoded  
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 CC sequence and the sequence of AAT04076 (encoding a 261 amino acid  
 CC protein) can be used to detect M.gallisepticum using probes based on  
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 CC complement of bases 893-919, 1908-1934 and the complement of bases  
 CC 2184-2210 of the sequence, and a probe based on nucleotides 718-41 of  
 CC AAT04076. The method using these sequences is faster i.e. is able to  
 CC detect M.gallisepticum, which causes avian chronic respiratory diseases,  
 CC within one day, from avian secretions, washings from the lung,  
 CC respiratory tract, nasal cavity, etc.  
 CC  
 XX  
 SQ Sequence 661 AA;  
 Query Match 34.6%; Score 803; DB 16; Length 661;  
 Best Local Similarity 41.4%; Fred. No. 4.2e-44;  
 Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;  
 Oy 51 PVGSTVIRLEFGCGSITKKDANPNNGQTOLEAARMEITD-----LINAKAMTLASIQ 103  
 Db 35 PPSGG-----NNNGCNTNPBGCGMMNAAKELADAKAALTITLINGETANLASYE 84  
 Oy 104 DYAKIEASLSAYSEAEFTVNNNNLNTLEQLMAKTLESAINQANTDKTTFDNEHPLVE 163  
 Db 85 DYAKIKSELTSAYETAKAVASAKTGATLENEVEAKTTLDAIKKASANDFDAQHGSLVE 144  
 Oy 164 AYKALKTLEQRATNLEGLSTAYNQIRNNVLDLYNKASSLITKTLDLINGTILDSBEI 223  
 Db 145 AYNNIKETLKEKENTLDSLANENVAIITNNLSLYEKANTITVTATLDEAT-GNIPVMSV 203  
 Oy 224 TVVRNINNTLSTINEOKTNADALNSPFKVIQNNESFPGTFTNAVOPSNYSFVAFS 283  
 Db 204 TQANODITNATSRILAMKQNDNLANSITKOSLVKNNITRDVANNQ-QPANISFVGF 262  
 Oy 284 ADVTPVNYKVARIVWNGDE-----PSSRIANTNSITDVSWIYSLAGNTKYQSPFN 337  
 Db 263 VNVDPNPNMFAQRKVMASENTPLATTPADATQQAASLTDVSWIYSLNGAEKATLSFRY 322  
 Oy 338 YG-BSYGLVFPFYKYKAAADANNVGLQKLNANVOVEF-----ATSTSAN--NTTA 387  
 Db 323 FGAERTALYPPFYKLVKTS--NVGLQYKLNAGDKQINFVQTPASGSDVAANEETMA 380  
 Oy 388 NP-----TPADEVKIVAKIVISGLRFGONTIELSVPTGEGMKNKAPMIGNITVLS 438  
 Db 381 SPAEQSAPYVDIKIAKVALSNLKFNSNTLIEFSVPTG-----KAPMIGMNTLSSNSE 435  
 Oy 439 -----NENN-----ADKIPGY 449  
 Db 436 VNKKKIYDLPFGNSFNENNPPAVLVLDLKG 467  
 RESULT 15  
 AAM11978  
 ID AAM11978 standard; Protein; 647 AA.  
 XX  
 AC AAM11978;  
 XX  
 DT 23-APR-1997 (first entry)  
 XX  
 DE Mycobacterium gallisepticum pmGAL.2 adhesin.

XX  
 KW Adhesin; pmGAL; mycoplasma; diagnosis; vaccine; vector;  
 KM respiratory disease; poultry; haemagglutinin.  
 XX  
 OS Mycoplasma gallisepticum strain S6.  
 XX  
 FH Key  
 FT Peptide  
 FT 1..25  
 FT /label= sig peptide  
 FT /note= "the signal peptide shows homology to  
 FT the pmGAL.3 signal peptide"  
 XX  
 PN CA2135330-A.  
 XX  
 PD 11-MAY-1995.  
 XX  
 PF 08-NOV-1994; 94CA-2135330.  
 XX  
 PR 20-APR-1994; 94US-0230312.  
 PR 10-NOV-1993; 93AU-0050593.  
 XX  
 PA (BROW/) BROWNING G F.  
 XX  
 PI Browning GF, Glew MD, Markham PF, Walker ID, Whitehear KG;  
 XX  
 DR WPI; 1995-241027/32.  
 XX  
 DR N-PSDB; AAT51531.  
 XX  
 PT New promoter region from a Mycoplasma gallisepticum adhesin gene -  
 PT useful when coupled to foreign antigen gene, for prodn. of  
 PT multivalent live vaccines, also new probes for detecting Mycoplasma  
 PT and manipulating its genome  
 XX  
 PS Disclosure; Fig 3; 81pp; English.  
 XX  
 CC Adhesin pmGAL.2 (AAM11978) and adhesin pmGAL.3 fragment (AAM11979)  
 CC are products of gene sequences (see also AAT51531) isolated from  
 CC Mycoplasma gallisepticum. DNA constructs incorporating the  
 CC promoter and/or signal sequences of the pmGA genes can be used  
 CC in the prodn. of multivalent live vaccines. The signal peptide  
 CC sequence is utilised where attachment of an exogenous antigen  
 CC gene to the mycoplasma cell membrane is required.  
 CC  
 XX  
 SQ Sequence 647 AA;  
 Query Match 30.6%; Score 711.5; DB 16; Length 647;  
 Best Local Similarity 39.0%; Fred. No. 3.8e-38;  
 Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;  
 Oy 53 VGSTVIRLEFGCGMSITKK-----DANPNNGQ-----TOLEAARMEITD 90  
 Db 15 IGSFVMLAASCTTPTNPPTNPFPSPGCGMNGGTNPBGCGMMNAASQELAAARMGLTT 74  
 Oy 91 LINAKAMTLASIDYAKIEASLSAYSEAEFTVNNNNLNTLEQLMAKTLESAINQANTD 150  
 Db 75 IFDSKANLGLVVDYKKTQNTLTKAYDAKTVLNNSSSTQKNAEAKTRLETAIRTAATS 134  
 Oy 151 KTTFDNEHPLNVEAYKALKTTLLEQRATNLEGLSTAYNQIRNNVLDLYNKASSLITKTL 210  
 Db 135 KQTFDEQHAELVAYKVKELKTTLSNETATLAPYADAQYAGIMHLSGLYDAKKAITTKLE 194  
 Oy 211 PLNGTLLDSNEITVRNINNTL--STINEOKTN-----ADALNSPFKVIQNNESFV 264  
 Db 195 PVEGDP-LTSAVMMANTKIYEAIKDEVLNQKNAKTLADSLSSIVKKTGYBE----- 249  
 Oy 265 GTFTNAVOPSNYSFVAFSADVTPV-----NYKYARRTW--NGDEPSSRIANT-- 312  
 Db 250 ---AHNKAQPANYSFVGKRWYTELDDQYFPMWDAERTIFNSDEP--RSISNTPAD 304  
 Oy 313 -----NSTDVSWIYSLAGNTKYQSPFNGBSTGYLFPFYKYKAAADANNVGLQKLN 367  
 Db 305 GQTMAPLSNYSWISYSLAGTAKYTLLEFTYGPSTGYLFPFYKLVNTSDQVGLLEYKLN 364

```

Oy      3 6 8  NNNVVQVFEASTSA-----NNTTANPFAVIDEIKYAKIVLSSLRFGONTIEISVPTGEG 422
Db      3 6 5  D-----ATRPSATFGSDQDTQMKKITPVANDINAKVTLAUNLFGSKNIEFSVEA--- 413

Oy      4 2 3  NNNNKVAPMIGNTLYLSENENNADKI 446
Db      4 1 4  --EKVSFMIGNTLYLSSSPNNMNKI 435

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:19:47 ; Search time 29 Seconds

(without alignments)  
665,302 Million cell updates/sec

Title: US-09-901-572A-3

Perfect score: 2324  
Sequence: 1 MHYERRNCIFFLVILYGTN.....SSNNENADKIPGRRPCTFL 456

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_Aa.\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1917	82.5	615	2	US-08-525-742-10 Sequence 10, Appl
2	1849	79.6	610	2	US-08-525-742-8 Sequence 8, Appl
3	1645	70.8	368	2	US-08-525-742-4 Sequence 4, Appl
4	1619	69.7	368	1	US-08-185-851A-4 Sequence 4, Appl
5	803	34.6	661	2	US-08-525-742-2 Sequence 2, Appl
6	314	13.5	865	1	US-07-803-633A-13 Sequence 13, Appl
7	181.5	7.8	10182	4	US-09-134-001C-3159 Sequence 3159, Ap
8	176.5	7.6	933	3	US-08-293-728-2 Sequence 2, Appl
9	176.5	7.6	933	3	US-09-421-868-2 Sequence 4, Appl
10	174.5	7.5	1041	1	US-08-220-151-4 Sequence 4, Appl
11	174.5	7.5	1041	1	US-08-413-118-4 Sequence 4, Appl
12	174.5	7.5	1041	1	US-08-473-446-4 Sequence 4, Appl
13	150.5	6.5	682	4	US-08-836-687B-37 Sequence 37, Appl
14	147.5	6.3	1073	3	US-09-541-782-6 Sequence 6, Appl
15	147.5	6.3	1073	4	US-09-723-820-6 Sequence 6, Appl
16	139.5	6.0	930	4	US-09-134-001C-5314 Sequence 5314, Ap
17	139	6.0	1004	4	US-09-268-347-30 Sequence 30, Appl
18	137.5	5.9	1002	4	US-09-268-347-24 Sequence 24, Appl
19	137.5	5.9	2137	4	US-09-134-001C-4463 Sequence 4463, Ap
20	137	5.9	1098	3	US-08-923-992A-8 Sequence 8, Appl
21	135	5.8	943	3	US-08-911-321-4 Sequence 4, Appl
22	134.5	5.8	1073	4	US-09-206-942-49 Sequence 49, Appl
23	134.5	5.8	1079	4	US-09-206-942-47 Sequence 47, Appl
24	133.5	5.7	1164	3	US-08-923-992A-2 Sequence 2, Appl
25	132	5.7	1104	3	US-08-923-992A-4 Sequence 22, Appl
26	129.5	5.6	903	3	US-08-804-439A-22 Sequence 22, Appl
27	129.5	5.6	903	3	US-08-720-229-22 Sequence 22, Appl

28	129.5	5.6	904	3	US-08-632-537-1 Sequence 1, Appl
29	129.5	5.6	904	5	PCT-US96-05316-1 Sequence 1, Appl
30	129.5	5.6	904	6	5244792-4 Sequence 4, Appl
31	129	5.6	1833	4	US-08-621-944A-4 Sequence 4, Appl
32	129	5.6	1833	4	US-08-945-567D-4 Sequence 4, Appl
33	129	5.6	1992	4	US-08-621-944A-3 Sequence 3, Appl
34	129	5.6	1992	4	US-08-945-567D-3 Sequence 3, Appl
35	129	5.6	2048	4	US-09-268-347-48 Sequence 48, Appl
36	128.5	5.5	903	1	US-08-220-151-8 Sequence 8, Appl
37	128.5	5.5	903	1	US-08-413-118-8 Sequence 8, Appl
38	128.5	5.5	903	3	US-08-473-446-8 Sequence 8, Appl
39	128.5	5.5	1095	4	US-09-206-942-45 Sequence 45, Appl
40	128.5	5.5	1101	4	US-09-206-942-43 Sequence 43, Appl
41	128	5.5	682	3	US-08-481-435-6 Sequence 6, Appl
42	127.5	5.5	1128	3	US-08-923-992A-6 Sequence 6, Appl
43	127.5	5.5	1164	3	US-08-923-992A-10 Sequence 10, Appl
44	127	5.5	891	1	US-08-042-747A-6 Sequence 6, Appl
45	126	5.4	666	3	US-08-961-083-2 Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-525-742-10  
Sequence 10, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Sasaki, Sakiko  
APPLICANT: Ohkawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEROFO  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSEE: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930

TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 615 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-10

Query Match 82.5%; Score 1917; DB 2; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1,5e-126;  
Matches 378; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSITKCDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 123  
DB 27 CMSITKCDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 86  
QY 124 NNUNATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEORATNLEGLS 183  
DB 87 NNUNATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEORATNLEGLA 146  
QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRRNINNTLSTINEOKTN 206  
QY 244 ADALNSFTKVIYONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 303  
DB 207 ADALNSFTKVIYONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFNSYGPSTGYLYPPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFNSYGPSTGYLYPPYKLVKADANNVGLQ 326  
QY 364 YKLNNGNVQVEFAFSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 423  
DB 327 YKLNNGNVQVEFAFSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386  
QY 424 MNKVAPMIGNIYLSNENNAADKI 446  
DB 387 MNKVAPMIGNIYLSNENNAADKI 409

## RESULT 2

US-08-525-742-8  
Sequence 8, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohnawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-8

Query Match 79.6%; Score 1849; DB 2; Length 610;  
Best Local Similarity 95.0%; Pred. No. 8.7e-122;  
Matches 363; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 64 CMSITKCDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 123  
DB 27 CMSITKCDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 86  
QY 124 NNUNATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEORATNLEGLS 183  
DB 87 NNUNATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEORATNLEGLS 146  
QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRRNINNTLSTINEOKTN 206  
QY 244 ADALNSFTKVIYONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 303  
DB 207 ADALNSFTKVIYONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFNSYGPSTGYLYPPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFNSYGPSTGYLYPPYKLVKADANNVGLQ 326  
QY 364 YKLNNGNVQVEFAFSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 423  
DB 327 YKLNNGNVQVEFAFSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386  
QY 424 MNKVAPMIGNIYLSNENNAADKI 445  
DB 387 MNKVAPMIGNIYLSNENNAADKI 408

## RESULT 3

US-08-525-742-4  
Sequence 4, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohnawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi



APPLICANT: Takahashi, Kiyochito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESS: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESS: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 05-074139

FILING DATE: 31-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 05-245625

FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP94/00541

FILING DATE: 31-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Mcleland, Le-Nhung

REGISTRATION NUMBER: 31,541

REFERENCE/DOCKET NUMBER: 950811

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-659-2930

TELEFAX: 202-8870357

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-525-742-4

Query Match 70.8%; Score 1645; DB 2; Length 368;  
Best Local Similarity 96.1%; Pred. No. 8.7e-108;  
Matches 324; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 64 CMSITKQDANRNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSA5EAEVTN 123  
DB 27 CMSITKQDANRNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSA5EAEVTN 86  
QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLS 183  
DB 87 NNLAATLEQLKMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLS 146  
QY 184 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLSDNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLSDNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFIKKVYQNNQESFVGTFTNANVQPSNYSFVAFSAADVTPVNYKYARRTVMNGDE 303  
DB 207 ADALNSFIKKVYQNNQESFVGTFTNANVQPSNYSFVAFSAADVTPVNYKYARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQPSFSNYSFGSTGYLFPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQPSFSNYSFGSTGYLFPYKLVKADANNVGLQ 326  
QY 364 YKLNGNVQVEFATSTSNANTTANPTAVDEIKVAK 400

DB 327 YKLNGNVQVEFATSTSNANTTANPTQQLMRKLK 363

## RESULT 4

US-08-185-851A-4

Sequence 4, Application US/08185851A

Patent No. 5489430

GENERAL INFORMATION:

APPLICANT: Saico, Shuji

APPLICANT: Ohkawa, Setsuko

APPLICANT: Fujisawa, Ayumi

APPLICANT: Iitani, Yoshikazu

APPLICANT: Aoyama, Shigemitsu

TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene

TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As

TITLE OF INVENTION: Well As Vaccines Utilizing the Same

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Armstrong, Westerman, Hattori, Mcleland &

ADDRESS: Naughton

STREET: 1725 K Street, N.W., Suite 1000

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0

SOFTWARE: ASCII from Word Perfect version 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,851A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Theresa M. Stevens-Smith

REGISTRATION NUMBER: 36,281

REFERENCE/DOCKET NUMBER: PO-8-A930918

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-659-2930

TELEFAX: 202-887-0357

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-185-851A-4

Query Match 69.7%; Score 1619; DB 1; Length 368;  
Best Local Similarity 95.5%; Pred. No. 5.8e-106;  
Matches 322; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 64 CMSITKQDANRNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSA5EAEVTN 123  
DB 27 CMSITKQDANRNNGQTQLEAARMELTDLINAKAMTLASLDQYAKIEASLSA5EAEVTN 86  
QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLS 183  
DB 87 NNLAATLEQLKMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLS 146  
QY 184 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLSDNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLSDNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFIKKVYQNNQESFVGTFTNANVQPSNYSFVAFSAADVTPVNYKYARRTVMNGDE 303  
DB 207 ADALNSFIKKVYQNNQESFVGTFTNANVQPSNYSFVAFSAADVTPVNYKYARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQPSFSNYSFGSTGYLFPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQPSFSNYSFGSTGYLFPYKLVKADANNVGLQ 326

QY 364 YKLNNGNVQVEFATSTANTTANPTAVDEIKYAK 400  
Db 327 YKLNNGNVQVEFATSTANTTANPTAVDEIKYAK 363

## RESULT 5

US-08-525-742-2  
Sequence 2, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Onkawa, Setuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohsawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iitani, Yoshihikazu  
APPLICANT: Aoyama, Shigemi  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESS: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-2

Query Match 34.6%; Score 803; DB 2; Length 661;  
Best Local Similarity 41.4%; Pred. No. 2e-48;  
Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;

QY 51 PVVGSVTVLRFSGMSITKDKANPNNGQQLLEAARELTD-----LINAAMTLASIQ 103  
Db 35 PPSGG-----NMNGGNTNTPSDGGQGMNMAAAKELADAKAALTTLTLINGETANLASYE 84  
QY 104 DYAKIEASISAVSEAEVTNNNLATLEQLKMAKNTLSAIAQANTDKTFPNEHNLVE 163

Db 85 DYAKIESELISAYTATAVAKGATLTNEVNEAKTTLDAAIKKAASAKNPFDAQHSIVE 144  
QY 164 AYKALKTTLEORATNEGLSSSTAYNQIRNULVDLKNKASLITKTLDPGLNGTILDSNEI 223  
Db 145 AYNNLKEKLEKEKTNLDSLANENYAAIRTNLNSLYEKANTIVATLTPAT-GNPEWMSV 203  
QY 224 TTVNRNINNTLSTNECKTNADLNSFEIKVIONNESFGFTNTNANVOPSNYSFVAFS 283  
Db 204 TOANQDITNATSRILAKQNDNLANSFIKOSLVKNULRVDAVNOE-QPANSFVGF 262  
QY 284 ADVTPVNYKVARRTVMNGDE-----PSSRLANTNSITDVSWIYSLAGNTKXQFSFSN 337  
Db 263 VNVDTPNMNFQQRKVMASENPPLATTPAEDATQQAASLTDVSWIYSLNGAEAKYTLFERY 322  
QY 338 YG-PTGTYLPPYKLVQAADANNTGLQKLNNGVQVEF-----ATISAN-NTTA 387  
Db 323 FGAEKTYLVFPPYKLVVTSQ--NVGLQKLVGSGTQKINFVQTPASGSSDVAANEESTMA 380  
QY 388 NP-----TPAVDEIKVAKIVLSGLRFGONTTELSPTEGEGNMKNVAPMIGNIYLS- 438  
Db 381 SPAMQSLPVDDIKIAKVALSNLKNSTIERSVPTG-----KAAPMIGNIYLSNSE 435  
QY 439 -----NENN-----ADKIPGY 449  
Db 436 VNKKIYDDLFGNSFNNENNPATAVVDLKG 467

## RESULT 6

US-07-803-633A-13  
Sequence 13, Application US/07803633A  
Patent No. 5369025  
GENERAL INFORMATION:  
APPLICANT: NAZERIAN, Keyvan  
APPLICANT: LEE, Lucy F.  
APPLICANT: YANAGIDA, No. 5369025oru  
APPLICANT: OGAWA, Ryohel  
APPLICANT: LI, Yi  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR  
TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 No. 5369025ch Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/803,633A  
FILING DATE: 19911210  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerold M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1644-103P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 865 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-803-633A-13

Query Match 13.5%; Score 314; DB 1; Length 865;  
Best Local Similarity 100.0%; Pred. No. 5.4e-14;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFPLIVLYGTNSPSTONTNSREVSVOLSEESTFYLCPPVGVSTVRL 60  
DB 1 MHYFRNCIFPLIVLYGTNSPSTONTNSREVSVOLSEESTFYLCPPVGVSTVRL 60

QY 61 E 61  
DB 61 E 61

RESULT 7  
US-09-134-001C-3159  
; Sequence 3159, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3159  
; LENGTH: 10182  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3159

Query Match 7.8%; Score 181.5; DB 4; Length 10182;  
Best Local Similarity 20.5%; Pred. No. 0.003;  
Matches 121; Conservative 70; Mismatches 202; Indels 197; Gaps 23;

QY 20 NSSPSTONTV-SREVSVOLSEESTFYLCPPVGVSTVIRLFCGMSITKKDANPNNGQ 78  
DB 3961 NSNPSEVNEAQLQCYEAEVQKLVNDIAHIL-----QKKN-N 4016

QY 79 TOLKARMLTDLIAKAMTLASLQD-----YAKIBASISAYSEAEVNNMLNATLEOLK 134  
DB 4017 SALVYAKKQLOOSVVDQPLTGMQDSINNFEAKKNEQSAIRKAEAYINNDAIAKQIS 4076

QY 135 MAK-----TNLESAINQANTDKTFEDNEHPNLVEAYKALKTTLE 173  
DB 4077 DEKSKVEQALHLNDAKQQLFADTTELQTAVOQLRRGDT-NKKKPRGIMAYNKAIQSLE 4135

QY 174 QRATVL-EGLSSTAYNQR-----NNLYDLNKKASLITKT---LDPLNGGLDLSNET 224  
DB 4136 TQITSKAKNNAVAIOKPIRTYQVEVNNAQOVNQLQOLTEALINOQLPSSNDALKAARLN 4195

QY 225 TVNRNINNTLSTINEOKTNADALNSFIKKVIONNEOSFVGFTN----- 269  
DB 4196 LENK-INQTVQTDGMDQOSIEAYQNA--KRAVAQNSNTALALINNGDDEQOITTETRV 4252

QY 270 -----ANVQPSNYSFVAFSADVTPV-----NYKARRTVMNGD 302  
DB 4253 NQQTNLTOALINGLTVNKEPLETAKTALQNNIDQVSTDMQOSVAVNNQGLQAKKEI 4312

QY 303 EBSSTILN-----TNSIT-----DYSWITSL-----AGTN 328  
DB 4313 NTINNVLANNPDVNAIKTKNKAERISNDLTOAKNNLOVDQPLEKIKROLDEIDQSTN 4372

QY 329 T--KYQSFSSNYGPS--TGILYFPYKLVKAAADANNGLQYKLANGVQOV----- 374  
DB 4373 TDMQTDSDVNDVNDLSAIIIEKGVNKLKRNPRVEQVKEVANAQOVIQDLQNAKPSL 4432

QY 375 -----EFATSTSANNTTANFTPAVDEI-KVAKIVLSGLRFG 409

DB 4433 VPDXTLOQEAKNRLNLSINQDTDMQTDQSLANNYNDKLAKARONLEKISKVL-----GG 4487  
QY 410 QNTI-ELSVPTGEGNNKVA-----PMIGIYLSNENNAKD 445  
DB 4488 OPTAEIRONTDEANNAKQALDTARQSLTLNREEPIYINHHNESHUNNAK 4537

RESULT 8  
US-08-293-728-2  
; Sequence 2, Application US/08293728D  
; Patent No. 6008341  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: McDevitt, Damien L.  
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
; FILE REFERENCE: 05344.105011  
; CURRENT APPLICATION NUMBER: US/08/293,728D  
; PRIOR FILING DATE: 1994-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 933  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-08-293-728-2

Query Match 7.6%; Score 176.5; DB 3; Length 933;  
Best Local Similarity 22.7%; Pred. No. 0.00026;  
Matches 112; Conservative 66; Mismatches 184; Indels 131; Gaps 23;

QY 2 HYFRNCIFPLIVL-----YGTNS-----PSTONTNSREVSVOLSEESTFYLCPP 52  
DB 9 HAIRKSGVAVSLVGLTIGFGLLSKEADSENSVQSDASMSKSNDSSSAARPKT 68

QY 53 VGSTVIRLFCGMSITKKDANPNNGQTL--EARMELTDLIAKAMTLASLQDYAKIEA 110  
DB 69 DDTNV-----SDTSTSNVNGEISVQNPQOETQOSSSTNATT----- 108

QY 111 SLSAYSEAEVNNMLNATLEQLMAKTNLESAINQANTDKTFEDNEHPNLVEAYKALKT 170  
DB 109 EETVTEGATTTTNOANTPATTOSSNTNAELVNO-TSNETTND--TNTVSSVNS-- 162

QY 171 TLEORATNLEGLSST-----AYNQIRNNLYDLNKK-----ASSLIT 206  
DB 163 --PQNSTNAENVSTQDSTATPSNNESAPQSDASKVDVNAVNTSAPRMAFSLAA 220

QY 207 KTLDPANGTLLDSNETTTNRNINNTLSTINEQ---KTN-ADALNSPIK---KVIQ 257  
DB 221 VAADAPAPAGTDI--TNQTLNVTVGIDSGTIVPHQAGYKLVNYPGSPNSAVKGDFTFKITV 279

QY 258 NNEOSFVQGTFTNNAVQPSNYSFVAFSADVTPVNVKVARRTVMNGDEPESRILAMTNSITD 317  
DB 280 PKEINLVGVSTAVP-----TMAQDQ-----VLANGVIDSD 312

QY 318 VSMVYSLAG--TNTKYQSFSSNYGPGTGILYFPYKLVKAAADANNGLQYKLANGVQOYEF 376  
DB 313 GNVYITFTDYNTVDVYKATLTMPA--YI-----DBENV-----KKTGVN----- 350

QY 377 ATSTSANNTANFTPAVDEIVAKI-----VLSGLRFGQNTIELSVPTGEGNNKAP 429  
DB 351 TLAIGIGSTTANKVLVDYEEKYGFYNLSIKGITDQIDKTNNYRQITIVNPSGDVNIAP 410

QY 430 MI-GNIYLSNEN 441  
DB 411 VLTGNLKPNTDSN 423

RESULT 9  
US-09-421-868-2  
; Sequence 2, Application US/09421868  
; Patent No. 6177084  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Timothy J.

APPLICANT: McDevitt, Damien L.  
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
FILE REFERENCE: 05344.105011  
CURRENT APPLICATION NUMBER: US/09/421,868  
CURRENT FILING DATE: 1999-10-19  
PRIOR APPLICATION NUMBER: 08/293,728  
PRIOR FILING DATE: 1994-08-22  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 2  
LENGTH: 933  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-421-868-2

Query Match 7.6%; Score 176.5; DB 3; Length 933;  
Best Local Similarity 22.7%; Pred. No. 0.00026;  
Matches 112; Conservative 66; Mismatches 184; Indels 131; Gaps 23;

QY 2 HYFRNCIFELIVL-----YGTNS-----PSTONTSEVSVQLSSESTFYLCP 52  
DB 9 HAIRKSGVAVLVGLIGFELSSKEADASENSVQDSASNSKSDSSVSAAPKT 68  
QY 53 VGSIVIRLEFGGMSITKKDAPNNGQOL--EAMMELTDLINAKAMTLASLDYAKIEA 110  
DB 69 DDTNV-----SDTKSSNTNNGEISVAQNPAGQETTSSTNATL----- 108  
QY 111 SLSSASVSEAEVYNNNLATLEQLKAKTNLESAINQANTDKTTPNEHNPVLEAYKALKT 170  
DB 109 EETPTGEGATTTTQANTPATTOSSNTAEELVNO--TSNETTPND--TNTVSSVNS--- 162  
QY 171 TLEORATNLEGLST-----AYNOIRNNLVVLNKK-----ASSLIT 206  
DB 153 --PQSTINAEVSTTQDSTETAPSNESAPQSTASNDVNOAVNTSAPMRAFSIAA 220  
QY 207 KTLDELNGTLLDSNEITTVNENINNTLSTINEQ---KTN--ADALNSFIK---KVIQ 257  
DB 221 VAADAPAGTDL--TNQLTNVTGIDSGTTPVPHQGVYLVANGFVPSNAVAGDTFKITV 279  
QY 258 NNEQSFVGTFTNANVQPSNYSFVAPSAVTPVNYKARRTVNGCPSPSRILANTNSITD 317  
DB 280 PKEILNNGVTSTAKPFP-----IMAGQO---VLANGVIDSD 312  
QY 318 VSMIYSLAG--TWTKYQSFSPSYGPGTGYLYFPYKLVKADANNVGLQYKLNANGVQVEF 376  
DB 313 GNVITTFIDYVNTKODVAKATLMPA--YI-----DPENV-----KKTGVN----- 350  
QY 377 ATSTSANNTANPTPAVDEIKVAKI-----VLSGLRFGQNTIELSVPTGEGNNKVA 429  
DB 351 TLATGISTTANKTVLVDEYKYGKFPYNLSIKGTIDQIDKTNNTYROTIVYVNSGDNIAP 410  
QY 430 MI-GNVIYSSNEN 441  
DB 411 VLTGMLKPNVDSN 423

RESULT 10  
US-08-220-151-4  
Sequence 4, Application US/08220151  
Patent No. 5529780  
GENERAL INFORMATION:  
APPLICANT: Paoleletti, Enzo  
APPLICANT: Limbach, Keith J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESS: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,151  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-220-151-4

Query Match 7.5%; Score 174.5; DB 1; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

QY 3 YFRNCIFELIVL-----FLIVLY----- 17  
DB 22 YFRNCIFELIVL-----FLIVLY----- 17  
QY 18 -----GTNSPSTONTSEVSVQLSSEB---STFYLCPVGSIVIRLE- 61  
DB 82 PRRTVAPPEVGGTTPKPTPTDPTDMSDMREALRAQIEANGSTYMCPPGSGTVRLRP 141  
QY 62 -----FGCMSTKKDAP-----NNGCTQ 80  
DB 142 PRACPDYKLGKPTFEGIAVIFKENIAPYKFNAYIYKNIIMTWGSSVAVTNNRYTDR 201  
QY 81 LEARMELTDLINAKAMTLASLDYAKIEASVSAEFTVNNNLATLEQLKAKTNL 140  
DB 202 VPYKQETIDILDRGMCL-----SKADYVRNNYQFTAPDRBEDREL 244  
QY 141 ESAINQANTDKTTPDNEHNPVLEAYKALKTLEORATNLEGLSSTAYNOIRNNLVDLNKK 200  
DB 245 PLKSKFETPQSR--GWH-----TYKFAVYVYKDVIVSTAMAGSSYTQITNRVADRVP 297  
QY 201 ASSLITKTLDELNGTLLDSNEITTVNENINNTLSTINEQNTND-----ALNSFIKV 255  
DB 298 PVSEITDTIDKFG-----KCSSKATVYRN--NHKVEAFNEEDKNPQDMPILASKYNSVSKA 351  
QY 256 IONNEQSFV-----GFTNANVOP-----SNYSFVAPSAVTPVNY----- 291  
DB 352 WHTNNEYTKGAAGFPHSGTSVNCIYEVVARSVYPTDSPAISTGVYIHNSPFGRLDG 411  
QY 292 KYARRVYVNGDE-----PSSRIANTNSITDVSW-----IYS 323  
DB 412 AHVEHTSYSDRPFQIGEGYPIDDLRLQLGAPVSRNLETPHVT--VAMNMTPKGRCVCT 470  
QY 324 LAG-----TNTKYQSFSPSYGPGTGYLYFPYKLVKADANNVGLQYKLNANGVQVEF 376  
DB 471 LAKWREIDEMLRDEYQSYR-----FTVVTISATPISNTS--QFEIV--RIRLGDC 517  
QY 377 ATSTSANNTANPTPAVDEIKVAKIIVLSGLRFGQNTIELSVPTGEGNNKVA 433  
DB 518 ATKEAAL-----AIDRIYKSKYKTHIQTG--TLETYIARG--GFLIARPMISN 563

RESULT 11

US-08-413-118-4  
Sequence 4, Application US/08413118  
Patent No. 568920  
GENERAL INFORMATION:  
APPLICANT: PAOLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,118  
FILING DATE: 29-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/220,151  
FILING DATE: 30-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-413-118-4

Query Match 7.5%; Score 174.5; DB 1; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

3 YFRNCIF-----FLVILY-----17  
22 YFRQCFPPSLGLIAATGSRHNGSSGLTRLARVYFLVGRPPVVGSGSTSEQ 81  
18 -----GTNSSPSTQNTVTSREVVSVQLSEB---STFLCPVPVGSIVIRLE- 61  
82 PRRVATPEVGSTPKPTDPTDMDBEALRASQIEANGPSTFWCPCPGSGSTVRLBP 141  
62 -----FGCMSITKKDNP-----NNGQTQ 80  
142 PRACPDYKLGKNGFTGIAVFKEKNIAPYKFKANIYKNIIMTTWSSGSSYAVATTIRYDR 201  
81 LEAAMELTDLINAKMTLASLQDYAKIEASLSAYSEATVNNNTLEQLMAKNTL 140  
202 VPVKQGEITDLIDRGMCLE-----SKADYRNNYQGTAPRDEPREL 244  
141 ESAINQANTDKTTFNEHPNLVEAYKALKITLEQRATNLEGLSTAYNQIRNNLVLDLYNK 200  
245 PLKPSKFTVPSR---GMH-----TYKFKATVYKQVIVSTAMAGSSYQITFRYADRVPI 297  
201 ASSLTITKPLNGTGLSDNSEITVNNINNTLSTINEOKTNAD-----ALSNFIRKV 255  
298 PVSEITDITDKFG-----KCSSKATVYRN-NHKVEAFNEKDPQDMPILASKYNSVGSKA 351

256 IONNEQSFV-----GFTNANVOP-----SNYSFVARSADVTVPVNY-----291  
352 WHTNNEYTKLGAGFPHSGTSVNCIYEVNARSVPYDPSAISTGVYIHMSPEFGALDG 411  
292 KYARRTYWNGDE-----PSSRLIANTNSTITDSW-----IYS 323  
412 AHVHTSYSSDRFOQIEGYPIDLTRLQLGAPVSRNPLEPHVT-VAMNTPKCGRVCT 470  
324 LAG-----TNTKYSPFSNYPGSGTGLVPEYKLVFAADANNVGLQYKLNNGVQVEF 376  
471 LAKREIDEMLRDEYQSSYR-----FTVKTISATPISNTS-OPEIN-RIRLQDC 517  
377 ATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTSGNNKVAPMIGN 433  
518 ATKRAAE-----AIDRIYKSKYKSTHIQTG--TLETYLRAG-GFLIARPMISM 563

RESULT 12  
US-08-473-446-4  
Sequence 4, Application US/08473446  
Patent No. 6017542  
GENERAL INFORMATION:  
APPLICANT: PAOLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,446  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/413,118  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-473-446-4

Query Match 7.5%; Score 174.5; DB 3; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

3 YFRNCIF-----FLVILY-----17  
22 YFRQCFPPSLGLIAATGSRHNGSSGLTRLARVYFLVGRPPVVGSGSTSEQ 81  
18 -----GTNSSPSTQNTVTSREVVSVQLSEB---STFLCPVPVGSIVIRLE- 61

Db 82 PRRTVATPEVGTGPKPTTDDPTDMSDMREALRASQIEANQSPFTFMCPPSSSTVVRLEP 141  
Qy 62 -----FGCMSTTKDANP-----NNGQO 80  
Db 142 PRACPDYKLGKNTGEGIAVIFKENIAPYKFKANIYYKNIIMTTWMSGSSAYATTTNRYTDR 201  
Qy 81 LEAARMELTDLINAKMTTLASLODYAKIEASLSAYSSEAEVTNNINATLEQJAKMAKTL 140  
Db 202 VPKVQVEITDLIRKMKL-----SKADYANNYQFTAFDRDEPREL 244  
Qy 141 ESAINQANTDKTTFDNEHNPVAYKALKTTLEQORATNLEGLSTAYNOIRNNLVLYNK 200  
Db 245 PLKPSKFTFQSR--GMH-----TYKFKATVYKQIVSTAMAGSSYTOITRYADRVPI 297  
Qy 201 ASLLIKTIDPLNGGLDLSNETTYNNRINNTLSTINECKTAD-----ALSNSTKYV 255  
Db 298 PVSEITDITDKFG-----KCSSKATVVR--NHKVEAFNDKPKQDPLIASKXNSVSKA 351  
Qy 256 IONNEOSFV-----GTFNANVOP-----SNYSFVAFSADVTPVNY-----291  
Db 352 WHTNTYTKIGAGFHHSGTNCIVEVDARSVPYPSFALSTGDVIMSPFGLRQD 411  
Qy 292 KXARTVMNGD-----PSSRIANTNSITDVS-----IYS 323  
Db 412 AHVEHTSYSSDPFOQIEGYPIDLDTROLGAPVSHFLETPHVT--VANNMTPKCGRYCT 470  
Qy 324 LAG-----TNTKQFSFNSYSPSTGYLYFPYKLVKADANNVGLQYTLNNGVQOUEF 376  
Db 471 LAKWEIDEMUDEYQSGYR-----FTYKTIASATFISNTS--QFEIN--RRLGDC 517  
Qy 377 ATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNMKVAPMIGN 433  
Db 518 ATRKEAAE-----AIDRIYKSKYSKTHIQG--TLERYLARG--GFLIAFPMISM 563

RESULT 13  
US-08-836-687B-37  
; Sequence 37, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 37  
; LENGTH: 682  
; TYPE: PR1  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-37

Query Match 6.5%; Score 150.5; DB 4; Length 682;  
Best Local Similarity 21.0%; Pred. No. 0.011;  
Matches 75; Conservative 60; Mismatches 131; Indels 91; Gaps 13;

Qy 120 ETVNNNNLATTLEQLMAKTNLESAINQANTDKTTFDNEHNPVAYKALKTTLEORATNL 179  
Db 38 ELINNSNNALISTGSTTDSINLGAOSPVAKSTRTB--LDVTGAKTLLQTSAAQK 93  
Qy 180 EGLSTAVNOIRNNLVLYNKASSLITKLDPLNGGLDLSNETTYNNRINNTLSTINE 239  
Db 94 EKVVLQGVQVSE-----FSKRDSTVTKENAVPGDELLEQSEVAVVSSIQK--KILDN 148  
Qy 240 QKTADALSNSFIKVIQNNNEOSFVGTNNANVOPSNYSFVAFSADVTPVNYKXARTV- 298  
Db 149 KKKRANFVTSSEFLIKEPSSNKSADAGVIDNS-----ASPLSYRRAKEVVS 193  
Qy 299 -----NNGDEPSSRIILANTNSITDVSWMYSLACTNTKQYFSSSNVQPS 341  
Db 194 LRQPLAKQKVEAQPILLISNSSEKKASVYTNSHDEWDYQW-----DMKY--VTNNGES 243

Qy 342 TGYLYFPYKLVKADANNVGLQYKLNNGVQOUEBATSAN-----NT 385  
Db 244 YA-LYQPSKFI-----SVGI-----IDSGIMEEHPDLSNLGNVFPKLVKGFQDNBEDE 293  
Qy 386 TANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNMKVAPMIGN--NITLSSNEN 441  
Db 294 TGNPSDIDV-----KMGHTEVAQOITANGNILLVAPQITVNIYRVGEN 338

RESULT 14  
US-09-541-782-6  
; Sequence 6, Application US/09541782  
; Patent No. 6284480  
; GENERAL INFORMATION:  
; APPLICANT: Nislow, Corey  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Antifungal Assay  
; FILE REFERENCE: 1015  
; CURRENT APPLICATION NUMBER: US/09/541,782  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1073  
; TYPE: PR1  
; ORGANISM: S.pombe  
US-09-541-782-6

Query Match 6.3%; Score 147.5; DB 3; Length 1073;  
Best Local Similarity 20.0%; Pred. No. 0.034;  
Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;

Qy 17 YGNSSPSTQNT-----SREVSQV-----QISEESTPYLCPPPVSGTVIR 59  
Db 542 YETNEAKITVAATLQYRESKEYIASLYEKLDRTERRNNKENNNF-----N 590  
Qy 60 LERCMSITK-----KDNPNNGO--TOLEARMELTDLINAKA-MTLASL-----ODYAKIE 109  
Db 591 LKFNLITLMSFHSFDETNNGYTTLLDNFASMEELNTHSNOLLISMRTIHEFQSLD 650  
Qy 110 ASLSAYSEAEVTNNINATLEQLMAKTNLESAINQANTDKTTFDNEHNP-----LVEA 164  
Db 651 EALOSASSCQVPRSSIDLIVSELKSKSLDLALHSLDIDISSQKLGNGISSEITEL 710  
Qy 165 YKALKTT--LEORATNLEGLSTAYNOIRNNLVLYNKASSLITKLDPLNGGLDLSN 221  
Db 711 QDKMKESYRQIVQELRSLVNLQHTHEBSQKELMVGANDIDALVKTCTSLNDADITLSD 770  
Qy 222 EIT-----TYNNRINNTLSTINQ--KTNAAL-----SNSFIK--VI 256  
Db 771 YISDQSKFESKQODLIANTIGKIVSNFLQONSLYTKADILSHLNDTNSNRKANEIM 830  
Qy 257 QNNEOSFVGTNNANVOPSNYSFVAFSADVTPVNYKXARTVNG-----DEPSSRIANT 312  
Db 831 NNNSEEL--RNA-----ASQAEIVGANKKEIQTIVENGSLDLSKKAHSNS 877  
Qy 313 NSITDVSWMYSLA-----GNTKYO-----PFSFVYSGSTGYLYFPYKLVKA 354  
Db 878 RSMWD-----HCLALAESQKQGVNLEVOITDLRLQKVEHSEBDNTKEHQOL--LDLLES 930  
Qy 355 ADANNVGL--QYTLNNGVQV-----FATSTANNNT-----ANPTPAVDEIKV 398  
Db 931 LVGNNDNLISITKPTHELKTIIDHVLKGTTSIANHNNELLGLGDSLCNLETTTIEDTSL 990  
Qy 399 AKIVLSGLRFGONTI-----ELSVPTGEGNNMKVAPMIGNIYLSNENNAKIPG 448  
Db 991 VKLETTGDTGPKRELPAATPSWTRDSSLIKETTNLNLNDSQKFFRERYTSSNQNEPDV-- 1048  
Qy 449 YRRP 452  
Db 1049 YDKP 1052

**RESULT 15**

US-09-723-820-6

; Sequence 6, Application US/09723820

; Patent No. 6468760

**GENERAL INFORMATION:**

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

APPLICANT: Beraud, Christophe

; TITLE OF INVENTION: Antifungal Assay

FILE REFERENCE: 1015

CURRENT APPLICATION NUMBER: US/09/723,820

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 09/541,782

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEO ID NOS: 10

SOFTWARE: FastSEO for Windows Version 4.0

SOFTWARE: FACE  
: SEO ID NO 6

; SEQ ID NO 6  
: LENGTH: 1073

LENGTH: 10/2  
TYPE: PPT

TYPE: PRI  
ORGANISM: S nombe

URGENT SM: 2  
116-09-723-830-6

Query Match 6.3%; Score 147.5; DB 4; Length 1073;

Best Local Similarity 20.0%; Pred. No. 0.034;

Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;

```

QY 17 YGTVNSPSTQAVT-----SREAVSV-----QISEESTFPLCPPVSTVIR 59
D 542 YEYNEAKTIVTATDLSOYRESKETIAYLEYLDTERTENKKNENNNF-----N 590
QY 60 LEFGCMSITK-----KDNPNNGO--TOLEAARWELTDLINAKA--MTIAYL---QDYAKIE 109
D 591 LKFNULTWLRASFHGSFTDETNGYFTLLNDPNASMBELLNTHSNQOLLIMTKTHERFOGLD 650
QY 110 ASLSGANSSEATVNNNTNATTEQLKMAKTULESAINQANTDKTTPDNENHN----LYEA 164
D 651 EALOSARSSCAVPNSSLDLIIVSELDKSDXNSLDLDALEHSLDODISMSQKLGNGISSELIEL 710
QY 165 YKALKTT---LEQRATNLEGSTVAYNIRNNVULVYKASLIKTIIDPLNGGTLUDSN 221
D 711 QKDMKESIROLVQELRSLYLNLOHTHEESQKELMVGVRNDIDALVKXTCTTSLNDADIIUSD 770
QY 222 EIT-----TVARNINNTLSTINEQ--KTNAADAL-----SNSEIKK---VI 256
D 771 YISQOKSKFESKQODLIANIGKIVSNFQEOQNESIYTAADIIHSHLNTNINSIRANEIM 830
QY 257 QNNEQSFGVTETNANVOPSNSYFAFASDVTVPVNYKARRYWNG---DEBSRILANT 312
D 831 NNRSEBFL--RNA-----ASQAIEIVGANKERIQKTVENGSOQLDCKSKAIHNS 877
QY 313 NSIIVDSWITSLA-----GTVTKYQ-----FSSNSYSPSTGYLYFPKVLKYA 354
D 878 RSMWD---HCLTALAESQOKGVNLEQVTLDRLLQKKEHSEBDNTRYKHQOL---LDLES 930
QY 355 ADANNVGL--QYKLVNGVQOVE---PATSTSANNT-----ANPPTAVDELKV 398
D 931 LVGNNDNLIDISIKTPHTELQKITDHYLVKGTISLANHTHIELGLGDESLCNLETTIEDTSL 990
QY 399 AKIVLSGLRFQONTI-----ELSVPTGEGNNMVAAPWIGNIYLSNENNADKIPG 448
D 991 VKLETTGOTPSKRELPATPSPWTRROSSLIKETTNNLMLDSQDKKRVRRTYTSSNQTNEPDV-- 1048
QY 449 YRRP 452
D 1049 YDKP 1052

```

Search completed: August 14, 2003, 10:24:53  
Job time : 32 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 00:24:09 ; Search time 416 Seconds

(without alignments)  
8474.678 Million cell updates/sec

Title: US-09-901-572A-1

Perfect score: 1306  
Sequence: 1 aaaaacatcgatcttcaat.....taaatcgttaccagtc 1306

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N.GeneSeq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	98.2	2346	15	AAQ7857
2	1264	96.8	1387	15	AAQ7854
3	1262.4	96.7	1387	14	AAQ53419
4	1260.8	96.5	1387	16	AAQ94711
5	1086.8	83.2	1935	15	AAQ7856
6	1024.2	78.4	1371	18	AAQ7859
7	1022.6	78.3	3261	18	AAQ7859
8	989.4	75.8	1015	16	AAQ7859

9	824.2	63.1	853	11	AAQ04687	DNA encoding TMG-1
10	801.8	61.4	853	10	AAQ2574	DNA sequence of TM
11	679.8	52.1	708	10	AAQ2568	DNA sequence of M1
12	678.2	51.9	708	11	AAQ04686	DNA encoding MG-1
13	216.6	16.6	2196	16	AAQ74075	M.gallisepticum DN
14	216.6	16.6	2369	15	AAQ7853	Mycoplasm gallise
15	205.6	15.7	2452	15	AAQ68670	PMGAI.3 Mycoplasma
16	205.6	15.7	2453	16	AAQ51536	Mycobacterium gall
17	183.8	14.1	2416	15	AAQ68671	PMGAI.4 Mycoplasma
18	183.2	14.0	2417	16	AAQ51537	Mycobacterium gall
19	174.8	13.4	2809	16	AAQ51531	Mycobacterium gall
20	173.2	13.3	2282	16	AAQ51535	Mycobacterium gall
21	173	13.2	2881	15	AAQ68668	PMGA 1.2 and parti
22	160.6	12.3	2004	16	AAQ51538	Mycobacterium gall
23	150.8	11.5	2337	15	AAQ68669	PMGAI.2 Mycoplasma
24	149.2	11.4	2577	15	AAQ68672	Partial PMGAI.5 an
25	136.4	10.4	702	11	AAQ2570	DNA sequence of M3
26	136.4	10.4	702	11	AAQ05650	DNA encoding MG-3
27	78.2	6.0	594	10	AAQ2571	DNA sequence of M7
28	78.2	6.0	594	11	AAQ05651	DNA encoding MG-4
29	73.4	5.6	4985	22	AAQ75107	Anopheles gambiae
30	73	5.6	6292	22	AAQ546735	Tumour suppressor
31	69.6	5.3	7892	24	ABK40056	Human chemically p
32	68.6	5.3	8136	24	ABK39957	Human chemically p
33	68.6	5.3	8136	24	ABK32555	Human immune syste
34	68.4	5.2	11836	22	AAQ54395	Chemically pretrea
35	68.4	5.2	11836	22	ABK28240	DNA transcription
36	68.2	5.2	8392	24	ABK33490	Human immune syste
37	68	5.2	12405	22	AAQ54330	Human immune syste
38	68	5.2	12405	24	AAQ54330	Chemically pretrea
39	67	5.2	12405	24	ABK28169	Human gene regulat
40	67	5.1	8076	24	ABK39955	DNA transcription
41	66.6	5.1	8056	25	ABZ10246	Human chemically p
42	66.4	5.1	6669	24	ABK2197	Haemolysate of cel
43	66.4	5.1	6669	24	ABK2197	Chemically treated
44	66.4	5.1	6669	24	ABK2197	Human polynucleoti
45	66.2	5.1	969	24	ABQ39490	Human immune syste

#### ALIGNMENTS

AAQ7857	standard; DNA; 2346 BP.
AAQ7857	25-MAR-2003 (updated)
AAQ7857	23-JUN-1995 (first entry)
AAQ7857	Mycoplasm gallisepticum antigen TM-67 coding sequence.
AAQ7857	recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
AAQ7857	Mycoplasm gallisepticum.
AAQ7857	Location/Qualifiers
AAQ7857	202..2049
AAQ7857	/*tag= a
AAQ7857	/product= Mycoplasma _antigen
AAQ7857	202..2046
AAQ7857	/*tag= b
AAQ7857	/codon= seq:TGA,aa:Trp
AAQ7857	WO9423019-A1.
AAQ7857	13-OCT-1994.
AAQ7857	31-MAR-1994; 94WO-JP00541.
AAQ7857	31-MAR-1993; 93JP-0074139.
AAQ7857	30-SEP-1993; 93JP-0245625.

501 ATTATGGACACAGTGCACCAACCTTGAAGTTAGCTTCAACGCTTAATACAGATT 660

FD 13-OCT-1994.  
XX  
PF 31-MAR-1994; 94WO-JP00541.  
XX

FD 13-OCT-1994.  
XX  
PF 31-MAR-1994; 94WO-JP00541.  
XX

```

PR 31-MAR-1993; 93JP-0074139.
PR 30-SEP-1993; 93JP-0245625.
XX (JAPG) NIPPON ZEON KK.
PA (SHIO) SHIONOGI & CO LTD.
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohnawa I;
P1 Saeki S, Saito S, Takahashi K;
XX WPI: 1994-333181/41.
DR P-RSD8; AAR63227.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
PT exhibiting antigenicity of mycoplasma, useful for the production
PS of a live vaccine
XX
XX Claim 4; Page 71-74; 123pp; Japanese.
XX
XX The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding
CC for the 40kD antigen of Mycoplasma gallisepticum under the control
CC of a synthetic promoter. A 1300 bp restriction fragment containing
CC the promoter-ORF sequence was excised and was used in the
CC construction of plasmid pM7929-R2. This in turn was involved in the
CC construction of a recombinant avipox virus vector comprising the
CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
CC Newcastle Disease Virus haemagglutinin neuraminidase and FPV
CC sequences. The recombinant avipox virus is useful as a live vaccine
CC to protect against infection by Mycoplasma gallisepticum.
CC (Updated on 25-MAR-2003 to correct FN field.)
XX
XX Sequence 1387 BP; 531 A; 220 C; 203 G; 427 T; 6 other;
SQ
Query Match 96.8%; Score 1264; DB 15; Length 1387;
Best Local Similarity 98.3%; Pred. No. 5, 4e-196;
Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 1 AAAAAATCAGATGTTATATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60
DB 1 AAAAAATCAGATGTTATATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60
QY 61 AATCTTAATAATAAGCCGTTAAATTAACATAAAAAATTAATAAATGTTTCTTATC 120
DB 61 AATCTTAATAATAAGCCGTTAAATTAACATAAAAAATTAATAAATGTTTCTTATC 120
QY 121 AACCAAAATCTCTAGTATTAAGCGTTATTTATTTATTTTATTTAGCATCTTTAAAGT 180
DB 121 AACCAAAATCTCTAGTATTAAGCGTTATTTATTTATTTTATTTAGCATCTTTAAAGT 180
QY 181 ATAAATATATCTTATATTTCTATGATGAATAAGAAAGATCATCTTAAAGACTATTAAGTTG 240
DB 181 ATAAATATATCTTATATTTCTATGATGAATAAGAAAGATCATCTTAAAGACTATTAAGTTG 240
QY 241 TTAGGTACACATCTCTTTCTTATGCAATGGAATTTTACGCTGATCTTATTAATAAAA 300
DB 241 TTAGGTACACATCTCTTTCTTATGCAATGGAATTTTACGCTGATCTTATTAATAAAA 300
QY 301 GATGCAAAACCAAAATATGCGCAAAACCAATTAAGAGAGCGGAGATGATTAACAGAT 360
DB 301 GATGCAAAACCAAAATATGCGCAAAACCAATTAAGAGAGCGGAGATGATTAACAGAT 360
QY 361 CTAATCAATGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CTAATCAATGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 AGTTTATCATCTGCTTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 AGTTTATCATCTGCTTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GAAACAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 GAAACAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 AAAACGACTTTGATATGAACACCAAAATTTAGTTGAAGCATACAAAGCATTAACAAAC 600

```

```

DB 541 AAAACGACTTTGATATGAACATCAAAATTTAGTTGAAGCATTAACAAAGCATTAACAAAC 600
QY 601 ACTTTAAGACACGCTGATCACTTAACCTTGAAGGTTTGCATCACTGCTTATTAACAAAT 660
DB 601 ACTTTAAGACACGCTGATCACTTAACCTTGAAGGTTTGCATCACTGCTTATTAACAAAT 660
QY 661 CGCAATTAATTTAGTGAATCTTATACAAATTAAGCTAGTATTAATTAACATAACAGTAT 720
DB 661 CGTATATATTTAGTGAATCTTATACAAATTAAGCTAGTATTAATTAACATAACAGTAT 720
QY 721 CCCTAAATGGGGAGAGCTTTTATGATTTATGATGATTAATCACTTAATCGAATATT 780
DB 721 CCCTAAATGGGGAGAGCTTTTATGATTTATGATGATTAATCACTTAATCGAATATT 780
QY 781 AATAATGCTTATCACTTAATTAAGACAAAGACTAATGCTAGCTATTAATCTAATAGT 840
DB 781 AATAATGCTTATCACTTAATTAAGACAAAGACTAATGCTAGCTATTAATCTAATAGT 840
QY 841 TTTATTAATAAAGATTAATTAAGACAAAGACTAATGCTAGCTATTAATCTAATAGT 900
DB 841 TTTATTAATAAAGATTAATTAAGACAAAGACTAATGCTAGCTATTAATCTAATAGT 900
QY 901 AATGTTCACTTCAAACTACAGATTTTGTGCTTTAGTGTGATGATTAACACCGTCAAT 960
DB 901 AATGTTCACTTCAAACTACAGATTTTGTGCTTTAGTGTGATGATTAACACCGTCAAT 960
QY 961 TATTAATTAAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 TATTAATTAAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 AACAGAAATGATACAGATGTTTCTGATTTATGATTAAGCTGACAAACAGAGAG 1080
DB 1021 AACAGAAATGATACAGATGTTTCTGATTTATGATTAAGCTGACAAACAGAGAG 1080
QY 1081 TACCAATTTAGTTTACAGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 1140
DB 1081 TACCAATTTAGTTTACAGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 1140
QY 1141 TTGGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 TTGGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GTTCAACAAGTTGATTTGACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
DB 1201 GTTCAACAAGTTGATTTGACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
QY 1261 CCAGCAGTTGATGATTAAGTTGCTTAAGTTGCTTAAGTTGCTTAAGTTGCTTAAGTT 1306
DB 1261 CCAGCAGTTGATGATTAAGTTGCTTAAGTTGCTTAAGTTGCTTAAGTTGCTTAAGTT 1306

```

RESULT 3  
 AA053419  
 ID AA053419 standard; DNA; 1387 BP.  
 XX AC AA053419;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 16-JUN-1994 (first entry)  
 DE Mycoplasma gallisepticum antigen coding sequence.  
 XX  
 XX Vaccine; mycoplasma infection; poultry; fowl; ds.  
 OS Mycoplasma gallisepticum.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 202..1308  
 FT /tag=a  
 FT /product= 40kD\_antigen  
 FT /note= "ORF includes 2 NNN codons"  
 FT misc\_difference 51

FT FT /tag= b  
FT /note= "only partially legible in specification;  
FT misc\_difference 54 either C or G"  
FT /tag= c  
FT /note= "only partially legible in specification;  
FT misc\_difference 58 either C or G"  
FT /tag= d  
FT /note= "only partially legible in specification;  
FT misc\_difference 64 either C or G"  
FT /tag= e  
FT /note= "only partially legible in specification;  
FT either C or G"  
PN WO9324646-A1.  
XX 09-DEC-1993.  
PD  
XX 28-MAY-1993; 93WO-UP00715.  
XX  
XX 29-MAY-1992; 92JP-0138819.  
XX  
XX (JAFG ) NIPPON ZEON KK.  
XX (SHIO ) SHIONOGI & CO LTD.  
XX  
XX Aoyama S, Fujisawa A, Iricani Y, Ohkawa S, Satou S,  
XX WPI; 1993-405837/50.  
XX DR P-PSDB; AAR44493.  
XX  
XX Mycoplasma gallisepticum antigen and DNA coding for it - useful  
XX PT for vaccination of fowl against mycoplasma infections  
XX PS Claim 2; Page 23-26; 37pp; Japanese.  
XX  
XX The sequence coding for the 40kDa antigen was obtained by PCR  
XX CC amplification of M.gallisepticum genomic DNA. The antigen reacts  
XX CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used  
XX CC as a vaccine to protect fowl from M.gallisepticum infection.  
XX CC (Updated on 25-MAR-2003 to correct PN field.)  
XX SQ Sequence 1387 BP; 531 A; 216 C; 203 G; 427 T; 10 other;

Query Match 96.7%; Score 1262.4; DB 14; Length 1387;  
Best Local Similarity 98.0%; Pred. No. 9.9e-196;  
Matches 1280; Conservative 4; Mismatches 21; Indels 1; Gaps 1;

QY 1 AAAACATCAGATGTTATCTGATATCTTGTCTTAAAAAACAACAAATCTTCAACA 60  
DB 1 AAAACATCAGATGTTATCTGATATCTTGTCTTAAAAAACAACAAATCTTCAACA 60  
QY 61 AATCTTAATAATAAAGCCGTTAAATTAATAATAATAATAATAATAATAATAATA 120  
DB 61 AATCTTAATAATAAAGCCGTTAAATTAATAATAATAATAATAATAATAATAATA 120  
QY 121 AACCAAAATCTCTGTAATAAAGCTTATTTTATTTTATTTTATTTTATTTTAT 180  
DB 121 AACCAAAATCTCTGTAATAAAGCTTATTTTATTTTATTTTATTTTATTTTAT 180  
QY 181 AATAATATATCTTAATATCTTATGATATGAAGAAAGAAATCACTTAAAGACTTA 240  
DB 181 AATAATATATCTTAATATCTTATGATATGAAGAAAGAAATCACTTAAAGACTTA 240  
QY 241 TTAGGTACAAATCTTTCTTAGCATTTGGATTTTCTAGCTGTATGTCTTATTA 300  
DB 241 TTAGGTACAAATCTTTCTTAGCATTTGGATTTTCTAGCTGTATGTCTTATTA 300  
QY 301 GATGCAAAACCAAAATATGCGCAACCAATTAAGAGAGGCGCAATGAGTTAACA 360  
DB 301 GATGCAAAACCAAAATATGCGCAACCAATTAAGAGAGGCGCAATGAGTTAACA 360

QY 361 CTAATCAATGCTAANAGCATGATTTAGCTTCACTACAGACTATGCCAAGATTGAAGT 420  
DB 361 CTAATCAATGCTAANAGCATGATTTAGCTTCACTACAGACTATGCCAAGATTGAAGT 420  
QY 421 AGTTATATCTGCTTATATGAGAGCTGAACAGTAACATTAACCTTAATGCAACATTA 480  
DB 421 AGTTATATCTGCTTATATGAGAGCTGAACAGTAACATTAACCTTAATGCAACATTA 480  
QY 481 GAACAACTAAAAATGCTAAAACTAATTTAGATACAGCATCAACCAAGCTAATACGAT 540  
DB 481 GAACAACTAAAAATGCTAAAACTAATTTAGATACAGCATCAACCAAGCTAATACGAT 540  
QY 541 AAAAGACTTTTGTATATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAAC 600  
DB 541 AAAAGACTTTTGTATATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAAC 600  
QY 601 ACTTTAGAACACGCTGCTACTACCTTGAAGCTTTGTCATCACTGCTTATATCAATTT 660  
DB 601 ACTTTAGAACACGCTGCTACTACCTTGAAGCTTTGTCATCACTGCTTATATCAATTT 660  
QY 661 CGCAATTAATTTAGTGAATCTATACAAATTAAGCTAGTATTTAATTAATCACTAATAT 720  
DB 661 CGTATATATTTAGTGAATCTATACAAATTAAGCTAGTATTTAATTAATCACTAATAT 720  
QY 721 CCACCTAAATGGGGGAGAGCTTTTATGATTTCTAATGAGATTTACTACAGTTAATCGGAT 780  
DB 721 CCACCTAAATGGGGGAGAGCTTTTATGATTTCTAATGAGATTTACTACAGTTAATCGGAT 780  
QY 781 AATAATAGTTATCACTAATTAATGAACAAAGCACTAATGCTATGCTATATATAGT 840  
DB 781 AATAATAGTTATCACTAATTAATGAACAAAGCACTAATGCTATGCTATATATAGT 840  
QY 841 TTTATTTAAAAAGATTTCAAAATTAATGAACAAAGCTTTTGAAGCACTTTTACAAAGCT 900  
DB 841 TTTATTTAAAAAGATTTCAAAATTAATGAACAAAGCTTTTGAAGCACTTTTACAAAGCT 900  
QY 901 AATGTTCAACCTTCAAACTAGTTTGTGCTTTGCTGCTGTAACACCGGTCAT 960  
DB 901 AATGTTCAACCTTCAAACTAGTTTGTGCTTTGCTGCTGTAACACCGGTCAT 960  
QY 961 TATTAATATGCAAGAGAGCGTTGGAATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 961 TATTAATATGCAAGAGAGCGTTGGAATGATGATGATGATGATGATGATGATGATGAT 1020  
QY 1021 AACGCAATATGATCAAGATGTTTCTGATTTATTAATGTTAGTGAACAAACGAG 1080  
DB 1021 AACGCAATATGATCAAGATGTTTCTGATTTATTAATGTTAGTGAACAAACGAG 1080  
QY 1081 TACCAATTTAGTTTACCACTATGTCATGCAACCTGTTATTTATTTTCCCTATAG 1140  
DB 1081 TACCAATTTAGTTTACCACTATGTCATGCAACCTGTTATTTATTTTCCCTATAG 1140  
QY 1141 TTGTTTAAAGAGCTGATGCTAATTAACGTTGATTAATTAATTAATTAATTAAT 1200  
DB 1141 TTGTTTAAAGAGCTGATGCTAATTAACGTTGATTAATTAATTAATTAATTAAT 1200  
QY 1201 GTTCAACAAAGTTGATTTGCACTTCACTAGTGAAGTAATTAATTAATTAATTAAT 1260  
DB 1201 GTTCAACAAAGTTGATTTGCACTTCACTAGTGAAGTAATTAATTAATTAATTAAT 1260  
QY 1261 CCAAGAGTTGATGATTAAGTTGCTAATTAATGTTTATCAAGTT 1306  
DB 1261 CCAAGAGTTGATGATTAAGTTGCTAATTAATGTTTATCAAGTT 1306

RESULT 4  
AA094711  
ID AA094711 standard; DNA; 1387 BP.  
XX AA094711;  
AC  
XX  
DT 25-MAR-2003 (updated)  
DT 12-MAR-1996 (first entry)

XX XX Mycoplasma gallisepticum antigenic protein TTM-1 gene.  
 XX XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1, ds.  
 KW XX  
 XX XX Mycoplasma gallisepticum.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 202..1308  
 FT /\*tag= a  
 FT  
 XX JP07133295-A.  
 XX  
 XX 23-MAY-1995.  
 PD  
 XX 27-AUG-1993; 93JP-0213102.  
 PF  
 XX 27-AUG-1993; 93JP-0213102.  
 PR  
 XX (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 XX MPI: 1995-220782/29.  
 DR P-PSDB; AAR76955.  
 XX  
 XX A new antigenic protein which reacts with Mycoplasma gallisepticum -  
 PT is useful in a component vaccine for use against poultry infected  
 PT with M. gallisepticum.  
 XX  
 PS Claim 6; Figs 5-6; 33pp; Japanese.  
 XX  
 XX AA094711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein  
 CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum  
 CC infectious diseases in poultry, and as a diagnostic agent for  
 CC M. gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 XX Sequence 1387 BP; 531 A; 222 C; 201 G; 427 T; 6 other:  
 SQ  
 Query Match 96.5%; Score 1260.8; DB 16; Length 1387;  
 .Best Local Similarity 98.2%; Pred. No. 1.8e-195;  
 Matches 1282; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

Db 421 AGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAATGCAACATA 480  
 Qy 481 GAACAACTAAAAATGGCTAAAACTAATTTAGATCAGCATCAACCAAGTAATACGAT 540  
 Db 481 GAACAACTAAAAATGGCTAAAACTAATTTAGATCAGCATCAACCAAGTAATACGAT 540  
 Qy 541 AAAACGACTTTGATTAATGAACACCCAAATTTAGTGAAGCATACAAAGCACTAAAAAC 600  
 Db 541 AAAACGACTTTGATTAATGAACATCCAAATTTAGTCAACCATTAACCAAGCACTAAAAAC 600  
 Qy 601 ACTTTAGAACACAGCTGCTAATCACTTGAAGGTTTGCATCACTACTGCTTAATCAAAATT 660  
 Db 601 ACTTTAGAACACAGCTGCTAATCACTTGAAGGTTTGCATCACTACTGCTTAATCAAAATT 660  
 Qy 661 CGCAATATTTAGTGAATCTAATCAATTAAGCTAGTATTAAATCAATTAACCACTAGAT 720  
 Db 661 CGTAAATATTTAGTGAATCTAATCAATTAAGCTAGTATTAAATCAATTAACCACTAGAT 720  
 Qy 721 CCACATAATGGGGAGCGCTTTTATGATTTCTAATGAGATTACTACAGTTAATCGAATATT 780  
 Db 721 CCACATAATGGGGAGCGCTTTTATGATTTCTAATGAGATTACTACAGTTAATCGAATATT 780  
 Qy 781 AATATATCGTTATCACTAATTAATGAACAAAGCTAATGCTGATTCATTAATAGT 840  
 Db 781 AATATATCGTTATCACTAATTAATGAACAAAGCTAATGCTGATTCATTAATAGT 840  
 Qy 841 TTTATTAATAAAGATTCATAAATTAATGAACAAAGTTTGGAGGACTTTAGCAAAAGCT 900  
 Db 841 TTTATTAATAAAGATTCATAAATTAATGAACAAAGTTTGGAGGACTTTAGCAAAAGCT 900  
 Qy 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTATGCTGATGTAACACCCGTCAT 960  
 Db 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTATGCTGATGTAACACCCGTCAT 960  
 Qy 961 TATTAATATGCAAGAAGACCGTTTGAATGCTGATGAACCTTCAATGAATCTTGA 1020  
 Db 961 TATTAATATGCAAGAAGACCGTTTGAATGCTGATGAACCTTCAATGAATCTTGA 1020  
 Qy 1021 AACAGATATGATCACAGATGTTTCTGATTTATATGATTTAGCTGGAACCAACGAG 1080  
 Db 1021 AACAGATATGATCACAGATGTTTCTGATTTATATGATTTAGCTGGAACCAACGAG 1080  
 Qy 1081 TACCAATTTAGTTTACCAATATGTCATCACTGATTAATTAATTTCCCTTAATAG 1140  
 Db 1081 TACCAATTTAGTTTACCAATATGTCATCACTGATTAATTAATTTCCCTTAATAG 1140  
 Qy 1141 TTGTTAAAGACGCTGATGCTAATTAACGTTGATTAATCAATTAATTAATGAAT 1200  
 Db 1141 TTGTTAAAGACGCTGATGCTAATTAACGTTGATTAATCAATTAATTAATGAAT 1200  
 Qy 1201 GTTCAACAGTGAAGTTTGGCACTTCAACTAGTGAATTAATTAATTAATTAATCAACT 1260  
 Db 1201 GTTCAACAGTGAAGTTTGGCACTTCAACTAGTGAATTAATTAATTAATTAATCAACT 1260  
 Qy 1261 CCAGCAGTGTGATGATTAAGTTGCTAAATCGTTTATCAGGTT 1306  
 Db 1261 -CAGCAGTGTGATGATTAAGTTGCTAAATCGTTTATCAGGTT 1306

RESULT 5  
 AA077856  
 ID AA077856 standard; DNA; 1935 BP.  
 XX  
 XX AA077856;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 23-JUN-1995 (first entry)  
 XX  
 XX Mycoplasma gallisepticum antigen TM-66 coding sequence.  
 KW recombinant avipox virus; live vaccine; mycoplasma antigen; ds.  
 XX  
 XX Mycoplasma gallisepticum.

```

XX Key Location/Qualifiers
FH CDS 54..1886
FT /tag= a
FT /product= Mycoplasma_antigen
FT mat_peptide 54..1883
FT /tag= b
FT /codon= seq:TGA,aa:Trp
XX
XX MO9423019-A1.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94MO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohnawa I,
XX Saeki S, Saitos, Takahashi K;
XX WPI; 1994-333181/41.
XX P-PSDB; AAR63229.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX PUM-66 containing an open reading frame was sequenced (AA077856). The
XX ORF encodes an antigenic polypeptide. A recombinant avipox virus
XX comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1935 BP; 728 A; 314 C; 307 G; 586 T; 0 other;
XX
XX Query Match 83.2%; Score 1086.8; DB 15; Length 1935;
XX Best Local Similarity 96.7%; Pred. No. 2.5e-167;
XX Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
XX
XX 150 TTTATTTTATTTTATTTAGT CATCTTTTAAAGATATAATATCTTAATATTTCTATGAATA 208
XX 1 TTTATTTTATTTTGTGTAATCTTTTAAATATAATATTTTAAATATTTCTATGAATA 60
XX
XX 209 AGAAGAATCATCTTAAGACTATTAAGTTGTAGTAAACATCCCTTCTAGCATTTG 268
XX 61 AAAAAAGATCATCTTAAGACTATTAAGTTGTAGTAAACATCCCTTCTAGCATTTG 120
XX
XX 269 GGATTTCTAGCTGATGCTATCTATTAAGTAAAGATGCAACCAATATATGCGCAAAACC 328
XX 121 GGATTTCTAGCTGATGCTATCTATTAAGTAAAGATGCAACCAATATATGCGCAAAACC 180
XX
XX 329 AATTGAAGCAGCGCGAATGAGTTAAACAGATCTATCAATGCTAAAGCATGACATTAG 388
XX 181 AATTGAAGCAGCGCGAATGAGTTAAACAGATCTATCAATGCTAAAGCATGACATTAG 240
XX
XX 389 CTTCACTACAGACTATGATGCAAGATTTGAAGTTATTCATGCTTAATGTAAGTGTG 448
XX 241 CTTCACTACAGACTATGCAAGATTTGAAGTTATTCATGCTTAATGTAAGTGTG 300
XX
XX 449 AAACAGTTAACATTAACCTTAATGCAACATTAGAACTAAATATGCTTAAATCTAATT 508
XX 301 AAACAGTTAACATTAACCTTAATGCAACATTAGAACTAATAATATGCTTAAATCTAATT 360
XX
XX 509 TAGAATGAGCCATCAACCAAGTATATAGGATTAAGCACTTTTGATATATGAACCCCA 568
XX 361 TAGAATGAGCCATCAACCAAGTATATAGGATTAAGCACTTTTGATATATGAACCCCA 420

```

```

QY 569 ATTATGTTGAAGCATACAAAGCACTAAACCACTTTTGAACCAAGTCTACTAACTTTG 628
DB 421 ATTTAGTTGAAGCATACAAAGCACTAAACCACTTTTGAACCAAGTCTACTAACTTTG 480
QY 629 AAGGTTGTCATCAACCTGCTTATATGAATTTGCAATATTTTGTGATCTATACATA 688
DB 481 AAGGTTGTCATCAACCTGCTTATATCAATTTGCAATATTTTGTGATCTATACATA 540
QY 689 AAGCTAGTAGTTTATTAATCACTAAACACTAGATCCACTAAATGAGGGAACGCTTTAGATT 748
DB 541 AAGCTAGTAGTTTATTAATCACTAAACACTAGATCCACTAAATGAGGGAACGCTTTAGATT 600
QY 749 CTAATGAGATTTCTACAGTTATATGGAATATTAATATAGTTATCACTAATTAATGAC 808
DB 601 CTAATGAGATTTCTACAGCTATATAGATATTAATATTAATAGCTTATCACTAATTAATGAC 660
QY 809 AAAAGACTAATGCTGATGATCTATTAATAGTTTATTAATTAAGTATCAATAATATG 868
DB 661 AAAAGACTAATGCTGATGATCTATTAATAGTTTATTAATTAAGTATCAATAATATG 720
QY 869 AACAAAGTTTGTAGGACCTTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGTTTGTG 928
DB 721 AACAAAGTTTGTAGGATGTTTACAAACACTAATGTTCAACCTTCAAACTACAGTTTGTG 780
QY 929 TTGCTTTTATGCTGATGATGTAACACCCGCTCAATTTAAATATGCAAGAAAGCCGTTTGA 988
DB 781 TTGCTTTTATGCTGATGATGTAACACCTGTTATTAATTAATGCAAGAAAGCCGTTTGA 840
QY 989 ATGGATGTAACCTTCAAGTATGATCTTGTGCAACACAGAAATAGTATCAAGATGTTCTT 1048
DB 841 ATGGATGTAACCTTCAAGTATGATCTTGTGCAACACAGAAATAGTATGATGTTCTT 900
QY 1049 GGATTTATAGTTAGTCTGGAACAAACAGAGTACCAATTTAGTTTACCAATATGATGTC 1108
DB 901 GAATTTATAGTTAGTCTGGAACAAACAGAAATATCAATTTAGTTTACCAATATGATGTC 960
QY 1109 CATACACTGTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATCTATTAACG 1168
DB 961 CATACACTGTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATCTATTAACG 1020
QY 1169 TTGATTTACATACAAATTAATTAATGAATGTTTCAACAAAGTTGATGCTGCACTTCAA 1228
DB 1021 TTGATTTACATACAAATTAATTAATGAATGTTTCAACAAAGTTGATGCTGCACTTCAA 1080
QY 1229 CTAGTCAATATATATCTACAGCTATATCAACTCCAGCAGTTGATGATTAAGTTGCTA 1288
DB 1081 CTAGTCAATATATATCTACAGCTATATCAACTCCAGCAGTTGATGATTAAGTTGCTA 1140
QY 1289 AAATCGTTTATACAGTT 1306
DB 1141 AAATCGTTTATACAGTT 1158

```

RESULT 6  
 AAT96595  
 ID AAT96595 standard; DNA, 1371 BP.  
 XX AAT96595;  
 XX 15-JUL-1998 (first entry)  
 XX  
 XX Hybrid Marek's disease virus/M. gallisepticum gene 40 K-S.  
 XX  
 XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
 XX antigen; vaccine; poultry; ds.  
 XX  
 XX Chimeric - Marek's disease gammaherpesvirus.  
 XX Chimeric - Mycoplasma gallisepticum.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..1371  
 FT /tag= a

FT		/product= "protein 40 K-S"
FT	misc_feature	1..192
FT		/*tag= b
FT		/note= "derived from Marek's disease virus gb gene"
FT	misc_feature	193..1368
FT		/*tag= c
FT		/note= "derived from gene encoding M. gallisepticum antigen"
XX		
PX		
PI	Saito S, Tsuzaki Y, Yanagida N;	
XX		
DR	WPI; 1997-503046/46.	
XX	P-PSDB; AAM36050.	
PT	Fusion protein comprising herpes virus outer membrane protein and antigenic polypeptide - for prevention of infection by Mycoplasma gallisepticum, especially in poultry	
XX		
PS	Disclosure; Page 16-19; 51pp; Japanese.	
CC	This sequence encodes the chimeric protein 40 K-S which comprises a fragment of the Marek's disease virus outer membrane protein gb fused to an antigenic protein from Mycoplasma gallisepticum. The chimeric protein can be used in recombinant live vaccines for prevention of infection by Mycoplasma gallisepticum, especially as the outer membrane protein shows antigenicity in poultry.	
XX		
SO	Sequence 1371 BP; 491 A; 242 C; 232 G; 406 T; 0 other;	
Query Match	78.4%; Score 1024.2; DB 18; Length 1371;	
Best Local Similarity	99.7%; Pred. No. 3.4e-157;	
Matches 1026; Conservative	0; Mismatches 3; Indels 0; Gaps 0	
Dy	278 GCTGTATGTCATTACTAATAAAAGATGCACCAACCATAATATGGCCAAACCAATTAGAAG	337
Db	188 GCTGTATGTCATTACTAATAAAAGATGCACCAACCATAATATGGCCAAACCAATTAGAAG	247
Qy	338 CAGGCCGAATGGAAGTTAACAGATCTTAATCAATGCTTAAAGCGATGACATTAGCTTCACTAC	397
Db	248 CAGGCCGAATGGAAGTTAACAGATCTTAATCAATGCTTAAAGCGATGACATTAGCTTCACTAC	307
Qy	398 AAGACTATGCCAAGATTGAAGTAGTATTATCATCTGCTTAATAGTGAGCGTGAACAGTTA	457
Db	308 AAGACTATGCCAAGATTGAAGTAGTATTATCATCTGCTTAATAGTGAGCGTGAACAGTTA	367
Qy	458 ACAATTAACCTTAATGCAACATTAGAACCACTAAAAATGGCTAAAATAATTTAGATCAG	517
Db	368 ACAATTAACCTTAATGCAACATTAGAACCACTAAAAATGGCTAAAATAATTTAGATCAG	427
Qy	518 CCATTAACAAGCTTAATAGCGATTAAAGACATTTTGATTAATGAACACCCAATTTTAGTGG	577
Db	428 CCATTAACAAGCTTAATAGCGATTAAAGACATTTTGATTAATGAACACCCAATTTTAGTGG	487
Qy	578 AAGCATACAAAGCACTAATAAAACCACTTTAGAACCAAGTGTCTTAACCTTGAAGGTTGT	637
Db	488 AAGCATACAAAGCACTAATAAAACCACTTTAGAACCAAGTGTCTTAACCTTGAAGGTTGT	547
Qy	638 CATCAACTGCTTATTAATCAAAATTCGCCAATTAATTTAGTGATCTATACATAAAGCTAGTA	697
Db	548 CATCAACTGCTTATTAATCAAAATTCGCCAATTAATTTAGTGATCTATACATAAAGCTAGTA	607
Qy	698 GTTTAATAACTTAACACCTAGATCCACTTAATGGGGGAAAGCTTTTAGATTCTAATGAGA	757

Db	608	GTTTAATACAAACACTAGATCCACTAATATGGGGGAACGCTTTGATGATCTAATGGA	667
Qy	758	TTACTACAGTTAATCGGAATATTAAATAATGTTATCACTAATTAATGAACAAAGACTA	817
Db	668	TTACTACAGCTAATAAGAAATATATATATACGTTATCAACTAATTAATGAACAAAGACTA	727
Qy	818	ATGCTGATGCACTTATATATATGTTTATTTTAAAAAGATCAAAATTAATGAACAAAGTT	877
Db	728	ATGCTGATGCACTTATATATATGTTTATTTTAAAAAGATCAAAATTAATGAACAAAGTT	787
Qy	878	TTGTAGGAGACTTTTACAAACGCTAATGTTCAACCTCAAACTACAGTTTGTGCTTTTA	937
Db	788	TTGTAGGAGACTTTTACAAACGCTAATGTTCAACCTCAAACTACAGTTTGTGCTTTTA	847
Qy	938	GTGCTGATGTAACACCCGTCATTATAATAATATGCAAGAGAGCCGTTTGAAATGATG	997
Db	848	GTGCTGATGTAACACCCGTCATTATAATAATATGCAAGAGAGCCGTTTGAAATGATG	907
Qy	998	AACCTTCAGTGAATTTCTTGCAAAACAGAAATATATACAGAGTTTCTTGATTTATA	1057
Db	908	AACCTTCAGTGAATTTCTTGCAAAACAGAAATATATACAGAGTTTCTTGATTTATA	967
Qy	1058	GTTTAGCTGGAACAAACAGAAAGTACCAATTAGTTTAAAGCACTATGTCATCAACTG	1117
Db	968	GTTTAGCTGGAACAAACAGAAAGTACCAATTAGTTTAAAGCACTATGTCATCAACTG	1027
Qy	1118	GTTAATTATATATTCCTTATTAAGTTGGTTTAAAGCAGTGAATGCTAATACGTTGATTA	1177
Db	1028	GTTAATTATATATTCCTTATTAAGTTGGTTTAAAGCAGTGAATGCTAATACGTTGATTA	1087
Qy	1178	AATACAAATTAAATAAATGAATGTTTCAACAGTTGAGTTGGCACTTCAACTAGTGCA	1237
Db	1088	AATACAAATTAAATAAATGAATGTTTCAACAGTTGAGTTGGCACTTCAACTAGTGCA	1147
Qy	1238	ATAATPACTACAGCTAATCCAACTCCAGAGTTGATGATTAATAGTTGCTAAAAATCGTTT	1297
Db	1148	ATAATPACTACAGCTAATCCAACTCCAGAGTTGATGATTAATAGTTGCTAAAAATCGTTT	1207
Qy	1298	TATCAGGTT 1306	
Db	1208	TATCAGGTT 1216	
RESULT 7			
AAT96596			
ID	AAT96596	standard; DNA; 3261 BP.	
XX	XX	AAT96596;	
AC	XX		
DT	XX	15-JUN-1998 (first entry)	
DE	XX	Hybrid Marek's disease virus/M. gallisepticum gene 40 K-C.	
XX	XX		
KM	XX	Chimeric; Marek's disease virus; outer membrane protein; fusion protein;	
OS	XX	antigen; vaccine; poultry; ds.	
OS	XX	Chimeric - Marek's disease gammaherpesvirus.	
OS	XX	Chimeric - Mycoplasma gallisepticum.	
PH	Key	Location/Qualifiers	
FT	CDS	1..3261	
FT		/*tag= a	
FT		/product= "protein 40 K-C"	
FT	misc_feature	1..2016	
FT		/*tag= b	
FT		/note= "derived from Marek's disease virus gB gene"	
FT	misc_feature	2077..3258	
FT		/*tag= c	
FT		/note= "derived from gene encoding M. gallisepticum	
FT		antigen"	
XX	XX		
XX	XX	W09736924-A1.	

```

PD 09-OCT-1997.
XX
XX 28-MAR-1997; 97MO-JD01084.
XX
XX 29-MAR-1996; 96JP-0103548.
XX
XX (JAPG ) NIPPON ZEON KK.
XX
XX Saito S, Tsuzaki Y, Yanagida N;
XX
XX MPI; 1997-503046/46.
XX
XX P-PSDB; AAM36051.
XX
XX Fusion protein comprising herpes virus outer membrane protein and
XX PT antigenic polypeptide - for prevention of infection by Mycoplasma
XX PT gallisepticum, especially in poultry
XX
XX Disclosure; Page 22-30; 51pp; Japanese.
XX
XX This sequence encodes the chimeric protein 40 K-C which comprises a
XX CC fragment of the Marek's disease virus outer membrane protein 9B fused
XX CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
XX CC protein can be used in recombinant live vaccines for prevention of
XX CC infection by Mycoplasma gallisepticum, especially as the outer membrane
XX CC protein shows antigenicity in poultry.
XX
XX Sequence 3261 BP; 1099 A; 617 C; 648 G; 897 T; 0 other;
XX
XX Query Match 78.3%; Score 1022.6; DB 18; Length 3261;
XX Best Local Similarity 99.6%; Pred. No. 5.9e-157;
XX Matches 1025; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 278 GCTGTATGCTTATTAATAAAGATGCAACCCAAATATAGCCAAACCAATTGAAAG 337
DB |||||||
DB 2078 GCTGTATGCTTATTAATAAAGATGCAACCCAAATATAGCCAAACCAATTGAAAG 2137
QY 338 CAGCGCGAATGAGTGAACAGATCTATCATGCTTAAAGCGATGACATTAGCTTAC 397
DB |||||||
DB 2138 CAGCGCGAATGAGTGAACAGATCTATCATGCTTAAAGCGATGACATTAGCTTAC 2197
QY 398 AAGACTATGCCAAGATTGAAGCTTATTCATGCTTATATAGTGAAGCTGAACAGTTA 457
DB |||||||
DB 2198 AAGACTATGCCAAGATTGAAGCTTATTCATGCTTATATAGTGAAGCTGAACAGTTA 2257
QY 458 ACAATATACCTTATGCAACATTAAGAACTAATAATGGCTAATCTAATTTAGATCG 517
DB |||||||
DB 2258 ACAATATACCTTATGCAACATTAAGAACTAATAATGGCTAATCTAATTTAGATCG 2317
QY 518 CCATCAACCAAGCTATATGAGTAAACGACCTTTGATATGAAACCAACCAATTAGTTG 577
DB |||||||
DB 2318 CCATCAACCAAGCTATATGAGTAAACGACCTTTGATATGAAACCAACCAATTAGTTG 2377
QY 578 AAGCATACAAAGCATTAATAAAGCACTTTTGAACAACAGTCTTACCTTGAAGTTGT 637
DB |||||||
DB 2378 AAGCATACAAAGCATTAATAAAGCACTTTTGAACAACAGTCTTACCTTGAAGTTGT 2437
QY 638 CATCACTGCTTATATCAATTTGCAATTAATTTGATGATATATACATAAAGCTAGTA 697
DB |||||||
DB 2438 CATCACTGCTTATATCAATTTGCAATTAATTTGATGATATATACATAAAGCTAGTA 2497
QY 698 GTTATATATCTAATAACACTAGATCCACTAATATGGGGGAACGCTTTTATGATTTAG 757
DB |||||||
DB 2498 GTTATATATCTAATAACACTAGATCCACTAATATGGGGGAACGCTTTTATGATTTAG 2557
QY 758 TTACTACAGTTATATGGGAATTTAATATAGCTTATCAACTTATATGAAACAAAGACTA 817
DB |||||||
DB 2558 TTACTACAGTTATATGGGAATTTAATATAGCTTATCAACTTATATGAAACAAAGACTA 2617
QY 818 ATGCTGAGCATTTATCTATATAGTTTATTAATAAAGTGTGTAATAATATGAAACAAAGT 877
DB |||||||
DB 2618 ATGCTGAGCATTTATCTATATAGTTTATTAATAAAGTGTGTAATAATATGAAACAAAGT 2677
QY 878 TTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTA 937

```

```

DB 2678 TTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTA 2737
QY 938 GTGCTGATGTAAACCCGTCATTTAATAATATGCAAGAAGACCGTTTGAATGTGATG 997
DB |||||||
DB 2738 GTGCTGATGTAAACCCGTCATTTAATAATATGCAAGAAGACCGTTTGAATGTGATG 2797
QY 998 AACCTTCAAGTGAATTTCTTGCAAAACGAAATAGTATCAAGATGTTTCTTGATTTATA 1057
DB |||||||
DB 2798 AACCTTCAAGTGAATTTCTTGCAAAACGAAATAGTATCAAGATGTTTCTTGATTTATA 2857
QY 1058 GTTATGCTGAACAACGAAAGTACCAATTTAGTTTATAGCAACATAGTCCATCAACG 1117
DB |||||||
DB 2858 GTTATGCTGAACAACGAAAGTACCAATTTAGTTTATAGCAACATAGTCCATCAACG 2917
QY 1118 GTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATGCTAATACGTTGATTAC 1177
DB |||||||
DB 2918 GTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATGCTAATACGTTGATTAC 2977
QY 1178 AATACAAATTAATATATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCA 1237
DB |||||||
DB 2978 AATACAAATTAATATATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCA 3037
QY 1238 ATAATACTACAGCTTATGCTCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTT 1297
DB |||||||
DB 3038 ATAATACTACAGCTTATGCTCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTT 3097
QY 1298 TATCAGGTT 1306
DB |||||||
DB 3098 TATCAGGTT 3106

```

```

RESULT 8
AAT04076
ID AAT04076 standard; DNA; 1015 BP.
XX
XX AAT04076;
XX
AC 19-JUL-1996 (first entry)
XX
DT 19-JUL-1996
XX
DE M.gallisepticum DNA sequence II encodes 261 amino acid protein.
XX
KW Detection; probe; primer; PCR; amplification; secretion; lung;
XX avian chronic respiratory disease; respiratory tract; nasal cavity; de.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT CDS 202..987
FT ||| /*tag= a
FT |||
PM JP07236498-A.
XX
PD 12-SEP-1995.
XX
PF 25-FEB-1994; 94JP-0052764.
XX
PR 25-FEB-1994; 94JP-0052764.
XX
PA (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
DR MPI; 1995-347462/45.
XX P-PSDB; AAR79911.
XX
PT Detection of Mycoplasma gallisepticum - for the quick detection,
XX i.e. within one day, of avian chronic respiratory diseases
XX
PS Claim 3; Page 10-11; 11pp; Japanese.
XX
XX This is the nucleotide sequence of a fragment of the Mycoplasma
XX CC gallisepticum genome which codes for a 261 amino acid protein.
XX CC This sequence and the sequence of AAT04075 (encoding a 661 amino

```



CC acid protein) can be used to detect *M.gallisepticum* using probes  
CC based on nucleotides 1125-1648 and primers based on nucleotides  
CC 443-466, the complement of bases 893-919, 1908-1934 and the  
CC complement of bases 2184-2210 of AAT04075, and a probe based on  
CC nucleotides 718-41 of this sequence. The method using these sequences  
CC is faster i.e. is able to detect *M.gallisepticum*, which causes avian  
CC chronic respiratory diseases, within one day, from avian secretions,  
CC washings from the lung, respiratory tract, nasal cavity, etc.  
XX  
SQ Sequence 1015 BP, 400 A, 162 C, 138 G, 315 T, 0 other;

Db	841	TTTATTAAAAAGTGTATCGAAATATATGACAAAGATTTTGTAGGGACCTTTTACAAAGCT	900
Oy	901	AATGTTCAACCTTCAAACTACAGTTTGTGTTTGTAGTCTGATGTACACCCGTCAT	960
Db	901	AATGTTCAACCTTCAAACTACAGTTTGTGTTTGTAGTCTGATGTACACCCGTCAT	960
Oy	961	TATTAATATAGCAAGAGACCGCTTGGATGTGTGTAACCTTCAAGTAGAATTC	1015
Db	961	TATTAATATAGCAAGAGACCGCTTGGATGTGTGTAACCTTCAAGTAGAATTC	1015

Query Match	75.88;	Score 989.4;	DB 16;	Length 1015;
Post Local Simulation:	00.48;	Dist 14.1;		

RESULT 9 .

1 AAAAACAAGTTGTTATCTGATCTCTTCTTAAAAAACAAGATCTTTACAA 60  
 99: 30.4%; Prev. NO. 1.4e-131;  
 99: 0; Mismatches 16; Indels 0; Gaps 0

XX	10-MAR-2003 (updated)
DT	08-OCT-1990 (first entry)
XX	
DE	DNA encoding TMG-1 antigen.
XX	
KW	Mycoplasma gallisepticum; poultry; vaccine; ss.
XX	
OS	Mycoplasma gallisepticum.
XX	
PN	JF02111795-A.
XX	
PD	24-APR-1990.
XX	
PE	02-JUN-1989; 89JP-0136343.
XX	
PR	02-JUN-1989; 89JP-0136343.
XX	
PA	(JAPG ) NIPPON ZEON KK.
XX	
PA	(SHIO ) SHIONOGI KK.
XX	
DR	WPI, 1990-169109/22.
XX	
DR	P-PSDB; AAR05082.
XX	
PT	Diagnostic and vaccine for poultry mycoplasma serum - utilizes
PT	antigen protein of the disease and recombinant vector
PT	incorporated with its coding gene.
XX	
PS	
XX	Claim 5; Fig 2; 20pp; Japanese.
XX	
CC	The DNA can be inserted into an expression vector for the prodn. of
CC	polypeptide which elicits an antigen-antibody reaction with anti-
CC	mycoplasma gallisepticum poultry sera. It may also be ligated to
CC	other DNA to produce fusion proteins with an N-terminal bacterial
CC	enzyme sequence.
CC	See also AA004686 and AA005649-53.
CC	(Updated on 10-MAR-2003 to add missing OS field.)
XX	
XX	
90	Sequence 853 BP; 329 A; 138 C; 128 G; 258 T; 0 other;

Db	601	ACTTTAGAAACAACGTCCTACTAACCTTGAGGTTTAGCTTCAACGCTTTATATCAGATT	660
Qy	661	CGCATAATTTTATGTGATCTATACATATAAGCTAGTATTATATACTBAACACTAGAT	720
Db	661	CGTATAATTTTATGTGATCTATACATATAAGCTAGTATTATATACTBAACACTAGAT	720
Qy	721	CCACTAAATGGGGGAAACGTTTATGATTCTMAATGAGATTACTACAGATTAAATGGAAATAT	780
Db	721	CCACTAAATGGGGGAAATGCTTTTATGATTTCTAATGAGATTACTACAGATTAAATGGAAATAT	780
Qy	781	AATAATACGTTATCAACTATTATATGACAAAAAGACTAAATGCTGATGACTTATCTAATAGT	840
Db	781	AATAATACGTTATCAACTATTATATGACAAAAAGCTAAATGCTGATGACTTATCTAATAGT	840
Qy	841	TTTATATAAAGTGATTCAAATATATGACAAAGTTTTTGTATGGACCTTTTCAAAACGT	900

Query Match	63.1%	Score 824.2	DB 11	Length 853
Best Local Similarity	97.9%	Pred 7.7e-125		
Matches 835	Conservative 0	Mismatches 18	Indels 0	Gaps 0
Qy	163	TTAGTCATCTTTTAAAGATATAAATATATCTTAATATTCATGATTAAGAAAAAGATATC	2223	
Db	1	TTAGTCATCTTTTAAAGATATAAATATATCTTAATATTCATGATTAAGAAAAAGATATC	60	
Qy	223	TTAAACATTAAGTTGTTAGTCAACATCCTTTCTTACATTTGGATTTCTAAGCTGT	287	
Db	61	TTAAACATTAAGTTGTTAGTCAACATCCTTTCTTACATTTGGATTTCTAAGCTGT	120	
Qy	283	ATGTCATTACTAAAAAAGATGCAAAACCAATATATGCGCAAAACCAATTGAAGCAGCG	3424	
Db	121	ATGTCATTACTAAAAAAGATGCAAAACCAATATATGCGCAAAACCAATTGAAGCAGCG	1808	
Qy	343	CGAATGAGATTAACGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAAC	4020	

```

Db      181 CGAATGAGTTAATGATCTAATCAATGCTAAAGCAAGACATTAAGCTTCACTCAAGAC 240
Qy      403 TATGCCAAGATTGAAGCTAGTTATCATCTGTTATATAGTGAAGCTGAACAGTTAACAAT 462
Db      241 TATGCTTAAGATTGAAGCTAGTTATCATCTGTTATATAGTGAAGCTGAACAGTTAACAAT 300
Qy      463 AACCTTAATGCAACTAGTAAGCAACTAAATAATGGCTAAATCTAATTTAGATAGGCATC 522
Db      301 AACCTTAATGCAACTAGTAAGCAACTAAATAATGGCTAAATCTAATTTAGATAGGCATC 360
Qy      523 AACCAAGCTAATACGATTAATAACGACTTTGATATGAACACCCAAATTTAGTTGAAGA 582
Db      361 AACCAAGCTAATACGATTAATAACGACTTTGATATGAACATCAAAATTTAGTTGAAGA 420
Qy      583 TACAAAGCACTAAACCACTTTAGTAACAACGCTGCTACTTAACCTTGAAGGTTGTCATCA 642
Db      421 TACAAAGCACTAAACCACTTTAGTAACAACGCTGCTACTTAACCTTGAAGGTTGTCATCA 480
Qy      643 ACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 702
Db      481 ACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 540
Qy      703 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 762
Db      541 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 600
Qy      763 ACGATTAAATCGGAATTTAATAATAGTTATCAACTTAATGAACAAAGACTAATGCT 822
Db      601 ACGATTAAATCGGAATTTAATAATAGTTATCAACTTAATGAACAAAGACTAATGCT 660
Qy      823 GATGCAATTAATCAATAGTTTATTAATAAAGTATCAAAATATGAACAAAGTTTGTGA 882
Db      661 GATGCAATTAATCAATAGTTTATTAATAAAGTATCAAAATATGAACAAAGTTTGTGA 720
Qy      883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTAGAGTTTGTGCTTTAGTGT 942
Db      721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTAGAGTTTGTGCTTTAGTGT 780
Qy      943 GATGTAAACCGGTCAATTAATAATATGAAGAAGACCGTTTGAATGTGATGAACCT 1002
Db      781 GATGTAAACCGGTCAATTAATAATATGAAGAAGACCGTTTGAATGTGATGAACCT 840
Qy      1003 TCAAGTAGAATTC 1015
Db      841 TCAAGTAGAATTC 853

RESULT 10
AAN92574
ID      AAN92574 standard; DNA; 853 BP.
XX
AC      AAN92574;
XX
DT      25-MAR-2003 (updated)
DT      11-MAY-1990 (first entry)
XX
DE      DNA sequence of TM-1 encoding polypeptide TMG-1.
XX
KW      Mycoplasma gallisepticum; Poultry vaccine; ss;
XX
OS      Mycoplasma gallisepticum.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
XX      /tag a
XX
PN      EP345021-A.
PD      06-DEC-1989.
XX
PF      31-MAY-1989; 89EP-0305441.
PF      31-MAY-1989; 89EP-0305441.

```

```

XX      02-JUN-1989; 89JP-0136343.
PR
XX      (JAPG ) NIPPON ZEON KK.
XX      (SHIO ) SHIONOGI SEIYAKU KK.
PA      (JAPG ) NIPPON ZEON KK.
PA      (SHIO ) SHIONOGI SEIYAKU KK.
XX
PI      Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
PI      Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
DR      WPI: 1989-358393/49.
XX      P-PSDB: AAP93959.
XX
PT      New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PS      vaccines.
XX
XX      Disclosure; Fig.2; 31pp; English.
CC      This base sequence of TM-1 encodes the TMG-1 polypeptide which has the
CC      same sequence as that of a polypeptide expressed in Mycoplasma
CC      gallisepticum in nature. When inserted into a recombinant vector used to
CC      transform a host the antigen protein produced can be used as a vaccine to
CC      prevent and diagnose MG infection.
CC      (Updated on 25-MAR-2003 to correct PF field.)
CC      (Updated on 25-MAR-2003 to correct PA field.)
CC      (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 853 BP; 330 A; 139 C; 128 G; 256 T; 0 other;

```

```

Query Match      61.4%; Score 801.8; DB 10; Length 853;
Best Local Similarity 96.2%; Pred.No.3.2e-121;
Matches 821; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

```

Qy      163 TTAGTCATCTTTTAAAGATATATATATCTTAATATCTAATGATGAATAAGAAAGATCATC 222
Db      1 TTAGTCATCTTTTAAAGATATATATATCTTAATATCTAATGATGAATAAGAAAGATCATC 60
Qy      223 TTTAAGACTATTAGTTGTTAGTGAACAACATCCTTTCTTACATTTGGATTTCTAGCTGT 282
Db      61 TTTAAGACTATTAGTTGTTAGTGAACAACATCCTTTCTTACATTTGGATTTCTAGCTGT 120
Qy      283 ATGTCTTTTAAAGATATGCAACCCAAATTAATGCGCAACCCAAATTTAGAAAGCAGCG 342
Db      121 ATGTCTTTTAAAGATATGCAACCCAAATTAATGCGCAACCCAAATTTAGAAAGCAGCG 180
Qy      343 CGAATGAGTTAACAGATCTAATCAATGCTAAAGCGATGAACATTAGCTTCACTACAAGAC 402
Db      181 CGAATGAGTTAACAGATCTAATCAATGCTAAAGCGATGAACATTAGCTTCACTACAAGAC 240
Qy      403 TATGCCAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAACAGTTAACAAT 462
Db      241 TATGCCAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAACAGTTAACAAT 300
Qy      463 AACCTTAATGCAACTAGTAAGCAACTAAATAATGGCTAAATCTAATTTAGATAGGCATC 522
Db      301 AACCTTAATGCAACTAGTAAGCAACTAAATAATGGCTAAATCTAATTTAGATAGGCATC 360
Qy      523 AACCAAGCTAATACGATTAATAACGACTTTGATATGAACACCCAAATTTAGTTGAAGA 582
Db      361 AACCAAGCTAATACGATTAATAACGACTTTGATATGAACATCAAAATTTAGTTGAAGA 420
Qy      583 TACAAAGCACTAAACCACTTTAGTAACAACGCTGCTACTTAACCTTGAAGGTTGTCATCA 642
Db      421 TACAAAGCACTAAACCACTTTAGTAACAACGCTGCTACTTAACCTTGAAGGTTGTCATCA 480
Qy      643 ACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 702
Db      481 ACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 540
Qy      703 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 762
Db      541 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 600

```



```

PR 02-JUN-1989; 89JP-0136343.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI KK.
XX
DR MPI; 1990-169109/22.
DR P-PSDB; AAR05081.
XX
PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes
PT antigen protein of the disease and recombinant vector
PT incorporated with its coding gene.
XX
PS Claim 3; Fig 1a; 20pp; Japanese.
XX
CC The DNA can be inserted into an expression vector for the prodn. of
CC polypeptide which elicits an antigen-antibody reaction with anti-
CC mycoplasma gallisepticum poultry sera. It may also be ligated to
CC other DNA to produce fusion proteins with an N-terminal bacterial
CC enzyme sequence.
CC See also AA004687 and AA005649-53.
CC (Updated on 10-MAR-2003 to add missing OS field.)
CC
XX Sequence 708 BP; 279 A; 119 C; 107 G; 203 T; 0 other;
SQ
Query Match 51.9%; Score 678.2; DB 11; Length 708;
Best Local Similarity 97.5%; Pred. No. 3.2e-101;
Matches 689; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 280 TGTATGCTATTACTTAAAGAGATGCAACCCAAATATATGCGCAACCCCAATTAGACGA 339
DB 1 TGTATGCTATTACTTAAAGAGATGCAACCCAAATATATGCGCAACCCCAATTAGACGA 60
QY 340 GGGCAATGAGTAAACAGATCTATCATGCTTAAAGCATACATCTTACTTACATA 399
DB 61 GGGCAATGAGTAAACAGATCTATCATGCTTAAAGCATACATCTTACTTACATA 120
QY 400 GACTATGCGCAATGAGTAAACAGATCTATCATGCTTAAAGCATACATCTTACTTAC 459
DB 121 GACTATGCGCAATGAGTAAACAGATCTATCATGCTTAAAGCATACATCTTACTTAC 180
QY 460 AATAACCTTATGCAACATTAGAACCACTTAAATATGCTTAAATTTAGATCAGCC 519
DB 181 AATAACCTTATGCAACATTAGAACCACTTAAATATGCTTAAATTTAGATCAGCC 240
QY 520 ATCAACCAAGCTAATACGGAATTAACGACTTTGTAATGAACACCAATTTAGTGA 579
DB 241 ATCAACCAAGCTAATACGGAATTAACGACTTTGTAATGAACACCAATTTAGTGA 300
QY 580 GCATCAAAAGCACTTAAACCACTTAAACGACTTAAACCTTAAACCTTAAACCTTAA 639
DB 301 GCATCAAAAGCACTTAAACCACTTAAACGACTTAAACCTTAAACCTTAAACCTTAA 360
QY 640 TCAACTGCTTATTAATCAATTTGCAATATTTAGTGAATCTTAAATTAAGTATAGT 699
DB 361 TCAACTGCTTATTAATCAATTTGCAATATTTAGTGAATCTTAAATTAAGTATAGT 420
QY 700 TTAATTAACCTTAAACCACTTAAATGCGGGAAGCGCTTTGATTTCTTAATAGATT 759
DB 421 TTAATTAACCTTAAACCACTTAAATGCGGGAAGCGCTTTGATTTCTTAATAGATT 480
QY 760 ACTACAGTTAATCGGAATTTAATTAATGCTTAACTTAAATTAATGAACAAAGCTAAT 819
DB 481 ACTACAGTTAATCGGAATTTAATTAATGCTTAACTTAAATTAATGAACAAAGCTAAT 540
QY 820 GCTATGCTTATCTTAATTAATTTTAAATTAATGAATTAATGAATGAATGAATTT 879
DB 541 GCTATGCTTATCTTAATTAATTTTAAATTAATGAATTAATGAATGAATGAATTT 600
QY 880 GTAGGAGCTTTTAAACAGCTTAAATGCTTAACTTAACTTAACTTAACTTAACTTAA 939
DB 601 GTAGGAGCTTTTAAACAGCTTAAATGCTTAACTTAACTTAACTTAACTTAACTTAA 660
QY 940 GCTATGTAACACCGGTCAATTTAATTAATTAATGAACAGACCGTTTG 986

```

```

DB 661 GCTATGTAACACCGGTCAATTTAATTAATTAATGAACAGACCGTTTG 707
RESULT 13
AAT04075
ID AAT04075 standard; DNA; 2196 BP.
XX
XX AAT04075;
XX
AC 19-JUL-1996 (first entry)
XX
DE M.gallisepticum DNA sequence I encodes 661 amino acid protein.
XX
KW Detection; probe; primer; PCR; amplification; secretion; lung;
KW avian chronic respiratory disease; respiratory tract; nasal cavity; ds.
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT CDS 171..2156
FT FT /*tag= a
FT FT misc_difference 828..830
FT FT /*tag= b
FT FT misc_difference 978..980
FT FT /*tag= c
FT FT misc_difference 1083..1085
FT FT /*tag= d
FT FT misc_difference 1911..1913
FT FT /*tag= e
FT FT /*codon= seq: TGA, aa: Trp
XX
XX JP07236498-A.
XX
PD 12-SEP-1995.
XX
PF 25-FEB-1994; 94JP-0052764.
XX
PR 25-FEB-1994; 94JP-0052764.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI & CO LTD.
XX
XX MPI; 1995-347462/45.
XX
DR P-PSDB; AAR79910.
XX
PT Detection of Mycoplasma gallisepticum - for the quick detection,
PT i.e. within one day, of avian chronic respiratory diseases
XX
XX Claim 2; Page 7-10; 11pp; Japanese.
XX
XX This is the nucleotide sequence of a fragment of the Mycoplasma
XX gallisepticum genome which codes for a 661 amino acid protein.
XX This sequence and the sequence of AAT04076 (encoding a 261 amino
XX acid protein) can be used to detect M.gallisepticum using probes
XX based on nucleotides 1125-1648 and primers based on nucleotides
XX 449-466, the complement of bases 893-919, 1908-1934 and the
XX complement of bases 2184-2210 of this sequence, and a probe based
XX on nucleotides 718-41 of AAT04076. The method using these sequences
XX is faster i.e. is able to detect M.gallisepticum, which causes avian
XX chronic respiratory diseases, within one day, from avian secretions,
XX washings from the lung; respiratory tract, nasal cavity, etc.
XX
SQ Sequence 2196 BP; 800 A; 353 C; 376 G; 667 T; 0 other;
Query Match 16.6%; Score 216.6; DB 16; Length 2196;
Best Local Similarity 55.1%; Pred. No. 1.5e-26;
Matches 548; Conservative 0; Mismatches 414; Indels 33; Gaps 5;
QY 329 AATTAGAACGAGCGGATGAGTAAACAGATCTAATCAATGCTTAAAGCATTAAG 388

```

```

Db 352 AATTAGCAGAGCAGCAAAAGCTGCTTAACCTTTGATTAAGTGAAGCAATCTTG 411
Qy 389 CTTGACCTACAGACTATGCAAGATTGAAGCTAGTTTATCATCTGCTATATGAGCTG 448
Db 412 CGTCTATGAAAGATGCTATGATCAAAAGGAAATTAACATCAGCGTGAACAGCTA 471
Qy 449 AAACAGTTAACAATTAACCTTAATGCAATTAAGAACTAAATGCTAAATCTAATT 508
Db 472 AAGCAGTTTCAAGCTAAACTGGTGCACTTAAATGAGGTTAATAGGCAAAACTCAT 531
Qy 509 TAGAATCAGCATCAACCAAGCTAATACGATAAACCACTTTGATATGACCCCA 568
Db 532 TAGATGCTGCTATTAATAAAGCTGCTAGCTAAGAAATGATTTGATCAGACGGGT 591
Qy 569 ATTTAGTTGAAGCATACAAAGACTAAACCACTTTGAGAACAGTCTACTAACCCTG 628
Db 592 CACTAGTGGAAAGCATATACATCTTAAAGAAACGTTAAAGAAAGAAAACTAATTTAG 651
Qy 629 AAGGTTTGCATCAACTGCTTATATCAATTTGCAATTAATTTAGTGCATATACATA 688
Db 652 ATTCTCTGCAACCAAAATTAATGACGAAATCAGAACTTATATGTTTATATGAAA 711
Qy 689 AAGCTAGTATTTAATACTAAACATGATCCAATAATGGGGAAACGCTTTAGATT 748
Db 712 AAGCAATACTATGTTGTTACAGCTACTTATAGACCCGCTACTGAAATATTCCTGAAGTTA 771
Qy 749 CTATAGATATTAATCACTTAATTCGAAATTAATAATACGTTATCACTATATGAAAC 808
Db 772 TGAGTG--TAAACACAGCTAATCAAGATTAATTAAGCAACTTCAAGACTAATAGCTT 828
Qy 809 AAAAGACTAATGCTGATCATTAATTAATAGTTTATTAATAAAGTATCAAAATAATG 868
Db 829 GAAACAAATAGCTGATTAATTTAGTAAAGTTTATCAACAGCTTATAGTAAATA 888
Qy 869 AACAAAGTTTGAAGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTGG 928
Db 889 ATTGACATAGAGTTG--ATGTAGCAATTAATCAGAGCAACACAGCAATTAACAGTTTG 945
Qy 929 TTGCTTTAGTGCATGATTAACACCCGCTCAATTAATAATGCAAGAAAGACCGTTTGA 988
Db 946 TTGGTTTATGTTAATGTTGATTAATCTCTTAACGTAATTTTGGCGCAAGAAAGTTTGGG 1005
Qy 989 ATGGATGAACTCTCAAGTGAATTTTGCAAACAGCAATAGTA----- 1033
Db 1006 CCTCTGAAATATCTCTTAGCACTACACAGCTGAAGATGCAACACAGAGCTGCAT 1065
Qy 1034 ---TCACAGATGTTCTTGAATTAATTAATGCTGAACAAACAGAGTACCAATTTA 1090
Db 1066 CTTTAAACAGATGTTTCAATGAATCTATTAATGATGCTGAAGCTAATATACACATTAA 1125
Qy 1091 GTTTTGAACATATG--TCATCACTGCTATATTAATTTCCCTTATAGTTGTTA 1147
Db 1126 GCTTTCGTTACTTGGAGCTGAAAAAGAGCTTATTAATTTCCCTTATTAATTTAGTTA 1185
Qy 1148 AAGCAGCTGATGCTAATTAAGCTGATTAACATACAAATTAATAATGAATGTTCAAC 1207
Db 1186 AAAC-----TAGTATTAATGTTGTTTACATATTAATTAATGTTGTTTACTTAAC 1239
Qy 1208 AAGTTGATTTGCCACTTCAACTAGTGAATTAATTAATTAATTAATTAATTAATTA 1267
Db 1240 AATATTAATTTGTAACAACTCAGCTTCGAGTTCAAGGATGTTGCTGCTATATGAAG 1299
Qy 1268 TTGATGATTAATTAATGCTAATAATGCTTTATCA 1302
Db 1300 AAATATGCTAGTCTGCTGAAATGACGACGCA 1334

```

RESULT 14  
AAQ77853  
ID AAQ77853 standard; DNA; 2369 BP.  
XX  
AC AAQ77853;

```

XX 25-MAR-2003 (updated)
DT 23-JUN-1995 (first entry)
DE Mycoplasma gallisepticum antigen TM-81 coding sequence.
XX recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
KV Mycoplasma gallisepticum.
XX
OS
XX
FH Key Location/Qualifiers
FT CDS 171..2156
FT /tag= a
FT /product= antigenic_polypeptide
FT mat_peptide 171..2153
FT /tag= b
FT /codon= seq:TGA,aa:Trp
FT
FT
PN MO9423019-A1.
PD 13-OCT-1994.
XX
XX 31-MAR-1994; 94MO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohnawa I;
XX PI Saeki S, Saitos, Takahashi K;
XX DR WPI; 1994-333182/41.
XX
XX P-PSDB; AAR63226.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX PT exhibiting antigenicity of mycoplasma, useful for the production
XX PT of a live vaccine
XX
XX Claim 4; Page 61-65; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX CC PUN-81 containing an open reading frame was sequenced (AAQ77853). The
XX CC ORF encodes an antigenic polypeptide. A recombinant avipox virus
XX CC comprising the coding sequence can be used as a live vaccine to
XX CC protect against infection by Mycoplasma gallisepticum.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2369 BP; 859 A; 376 C; 401 G; 733 T; 0 other;

Query Match 16.6%; Score 216.6; DB 15; Length 2369;
Best Local Similarity 55.1%; Pred. No. 1.5e-26;
Matches 548; Conservative 0; Mismatches 414; Indels 33; Gaps 5;

Qy 329 AATTAGAAGCAGCGGAGTGAAGTTAACAATTAATCAATTAATGATGAAGCTG 388
Db 352 AATTAGCAGAGCAGCAAAAGCTGCTTAACCTTTGATTAAGTGAAGCAATCTTG 411
Qy 389 CTTGACCTACAGACTATGCAAGATTGAAGCTAGTTTATCATCTGCTATATGAGCTG 448
Db 412 CGTCTATGAAAGATGCTATGATCAAAAGGAAATTAACATCAGCGTGAACAGCTA 471
Qy 449 AAACAGTTAACAATTAACCTTAATGCAATTAAGAACTAAATGCTAAATCTAATT 508
Db 472 AAGCAGTTTCAAGCTAATAACTGGTGCACTTAAATGAGGTTAATAGGCAAAACTCAT 531
Qy 509 TAGAATCAGCATCAACCAAGCTAATACGATAAACCACTTTGATATGACCCCA 568
Db 532 TAGATGCTGCTATTAATAAAGCTGCTAGCTAAGAAATGATTTGATCAGACGGGT 591
Qy 569 ATTTAGTTGAAGCATACAAAGACTAAACCACTTTGAGAACAGTCTACTAACCCTG 628

```

Db	592	CACCTAGTGGAGCATATATAACATCTAAAGAAACGTTAAAAAGAAAGAAAACTAAATTG	651
Qy	629	AAGGTTTGCAATCACTGCTTAATATCAAAATTCGCAATAATTGATGAGATCTATACAA	688
Db	652	ATTCTCTTGCAACGAAAAATTATGACGAACTACGAATCTAATCTTAATAGTTTATGAAA	711
Qy	689	AAGCTAGTAGTTTAATTAATCTAAACACCTAGATCCATAATGGGGAGCGCTTTGATT	748
Db	712	AGCCAAATCTATTTGTTACAGCTACTTTGACCCCTGCTACTGSAATATTCCTGAAGTTA	771
Qy	749	CTAATGAGATTTCTACAGTTAATCGGAATATTATATACGTTATCACTATTATGAAC	808
Db	772	TGAGTG---TAAACAAGCTAAATCAAGATATTAATACTAATCACTTAAGCTAATGCTT	828
Qy	809	AAAAAGCTATATGCTGATGATTAATCTAATAGTTTATTAATAAAAGATTCAAATATATG	868
Db	829	GAAAAAATAATCTGTATTAATTATGCTACAGTTTATCAAAACAGCTTTATGTTAATAAAA	888
Qy	869	AACAAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTG	928
Db	889	ATTGACTCTAGAGTTG---ATGATGACAAATATACAGAGACAACAGCAAAATTACAGTTTG	945
Qy	929	TTGCTTTTAGTGTGATGTTAACAACCCGCTAATTTATTAATATGACAGAGACCGTTTGG	988
Db	946	TTGTTTTAGTGTATATGTTGATCTCTTACCTGAATTTTGGCAAGAAAAAGTTTGG	1000
Qy	989	ATGTGTAGTAACCTTCAAGTAGAATTTCTGCAAAACAGAAATAGTA-----	1030
Db	1006	CCTCGAAATAATCTCCTTTAGCACTACACACGCTGAAATGCAACACAACAAGCTGCAT	1060
Qy	1034	---TACACAGATGTTTCTTGAGATTATATGTTTACTGCTGAAACAAACAGAACTCAATTTA	1090
Db	1066	CCTTACACAGATGTTTCAATGAAATCTATAGTTTAAATGCTGAAGCTAATACACATTTAA	1120
Qy	1091	GTTTAGCAACTATGAG---TCCATCACTAGGTTATTTATTTCCCTTAATAGTTGGTA	1140
Db	1126	GCTTTCGTTACTTGTGAGCTGAAAAACAGCTTACTTAATTTCCCTTAATTAATTAAGTTA	1180
Qy	1148	AAGCAGCTGATGCTAATAACGTTGATTACAAATACAAATTAATAATGAAAAATGTTCAAC	1200
Db	1186	AAAC-----TAGTATATATGTTGTTTACATATATTAATGTTAATGGTGTGATACTAATAC	1230
Qy	1208	AAGTTGAGTTGCCACTTCACTAATGCAAAATATATCTACAGCTATCCACTCCAGCAG	1260
Db	1240	AAATTAATCTTGTATACAAATCCAGCTTCTGTTCAAGTATGTTGCTGCTAATGAAGAAG	1290
Qy	1268	TTGATGAGATTAAAGTTGCTAATAATCGTTTATCA	1302
Db	1300	AACTATGCTAGTCTGCTGTGAATGCACTCAGCA	1334
RESULT 15			
XX	AAQ68670	AAQ68670 standard, DNA; 2452 BP.	
XX	AAQ68670;		
XX	AC		
DT	28-FEB-1995	(first entry)	
XX	DE		
XX	PMGAL.3	Mycoplasma gallisepticum DNA.	
KM	pmga;	adhesin gene complex; hemagglutinin; conserved sequences;	
KW	primers;	probes; amplification; polymerase chain reaction;	
KM	specific;	detection; PCR; 1.2; 1.3; ss.	
XX	Myoplasma	gallisepticum.	
XX	OS		
XX	AU9350593-A.		
XX	PN		
XX	26-MAY-1994.		
PD			
XX	10-NOV-1993;	93AU-0050593.	
XX			

PR		10-NOV-1992;	92AU-0005744.
PA	(UTME ) UNIV MELBOURNE.		
PA	(BROW/) BROWNING G F.		
XX			
PI	Browning GF, Glew MD, Markham PF, Walker ID, Whithear KG;		
XX			
DR	WPI; 1994-209061/26.		
PT	Recombinant DNA constructs for Mycoplasma gallisepticum - for		
PT	diagnosis, treatment and prophylaxis of poultry respiratory		
PS	disorders.		
XX			
XX	Example 1; Fig 4; 51pp; English.		
CC	AA068670 shows the DNA sequence of pMGAL.3 derived from a 10 kb insert		
CC	from a Mycoplasma gallisepticum genomic DNA library, detected by probes		
CC	based on the T3 and C7 peptides. PMGAL.3 contains a (GAAn) consensus		
CC	sequence (n=4 to 13). Mycoplasma gallisepticum infection in poultry,		
CC	humans and other animals is of economic importance to many industries		
CC	and it is desirable to produce effective vaccines and probes for its		
CC	detection. The sequences and probes and vaccine vectors of the		
CC	invention can be used for the diagnosis and treatment of Mycoplasma		
CC	gallisepticum infection, and for prophylaxis.		
XX			
SQ	Sequence 2452 BP; 913 A; 404 C; 420 G; 715 T; 0 other;		
	Query Match	15.7%; Score 205.6; DB 15; Length 2452;	
	Best Local Similarity	53.5%; Pred. No. 8.8e-25;	
	Matches 492; Conservative 0; Mismatches 404; Indels 24; Gaps 2		
QY	323 AAACCAATTAGAAAGCAGCGAATGAGTTAGATTAAGATCTAATCATCTAAAGCGATGA 382		
Db	410 AACCAACATTAGAGCAGCTGTAGAAAACCGTAAACCGACTACTTGTTGTCAGAAAAATACCA 469		
QY	383 CATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTATGTTATCATCTGTTAAGT 442		
Db	470 ATGTGGCTTGATATGCTGATTAATGCACAAATTCAAAGTACTTAAAGCAGCTGTTAATGA 529		
QY	443 AAGCTGAACAGTTAACATTAACCTTTAATGCAATTTAGAACAATAAAATGGCTAATA 502		
Db	530 CAGCTAAATCTGCATCAGAAAATACAAAGCCGCACTTTTGAAAATCTAAGATCTGCATCA 589		
QY	503 CTAAATTGAATCAGCCATCAACCAAGCTAATACGATTAACGACTTTGATTAATGAAC 562		
Db	590 CTACACTACAAGCAGCTATTGATTAAGCTGCTAATGAATAACGTTATTGATAGTGTGA 649		
QY	563 ACCCAAAATTTAGTTGAAGCATACGAAGACATTAACCACTTTGAACAACAGTGTACTA 622		
Db	650 ATCAACCTTTAGTACTGTCATATATATCTTAAACATCACTAAATCTTAAATACTA 709		
QY	623 ACCTTGAAGGTTGTCTCAACTGCTTAATATCAAAATTCGCAATATTTAGTGATCTAT 682		
Db	710 GTTTAGAAAGGTTATCTGAAAATTAATATGTGTGTAATTAATAATCACCTAAGTAACTGT 769		
QY	683 ACAATAAAGCTAGTAGTTTAATACTAAACATAGATCCACTAATATGGGGAAAGCTTT 742		
Db	770 TTGATTAAGGTATGAGCAATTAACACTTAACCAATTAGATCTTAATGAGGTAAAGACCTA 829		
QY	743 TAGATTCTAATGAGATTACTACAGTTATCGAATTTAATAATATACGTTATCAACTATTA 802		
Db	830 CTTTAGGAAATGTAAGAAAGCTTAACAAGGATATTTAAATATGGCTTAATGCTGAATGCT 889		
QY	803 ATGAACAAAAGACTAATGCTGATGATCTATATCTAATAGTTTTTAATTAATAAATGATTCAA 862		
Db	890 TGAAGAAATGAAAAAGCAATGACAGACAAATTTAATGAATTTGAGAAAAACCCTGTTCTA 949		
QY	863 ATATATGAACAAGTTTGTATGAGGACTTTTCAACGCTAATAGTTCAACCTTGAACATACA 922		
Db	950 AAGAAAAATTAAATCAACAAGCGATACAGCTCAATTAATCAAGAAACAACGAAATATGAA 1009		
QY	923 GTTTTGTGCTTTTAGTGTGATGATTAAC-----ACCGTCAATTAATTAAT 967		

```

Db      1010 GTTTCAGCTTAAGTGTGATCTAAGTAAATTCGCAAAATCTAAGTAAAT 1069
Qy      968 ATGCAAGAAGGACCGTTGGAATGGTGAATGAACCTTCAAGTAGAA-----TTCTTG 1018
Db      1070 TTGCTCAAGAAGAGTTGAACTAGTGAATAATCAACAAACCGGAAAAACAGCTAGTTT 1129
Qy      1019 CAACACGAATAGTATCAAGATGTTTCTTGATTTATAGTTTACCTGGAACAAACACGA 1078
Db      1130 CCTCTCCTGTTTCTGCAACTGATGTTTCTTGATTTATAGTTTACCTGGAAGGAACTA 1189
Qy      1079 AGTACCAATTTAGTTTACCACTATGTCATCAACTGTTATTTATTTCCCTTATA 1138
Db      1190 AATATACATTAACCTTTAGATTAATGATCCGATTAATGCAATTTATATTTGCCCTATA 1249
Qy      1139 AGTGTAAAGACAGCTGATGCTAATAAGTTGATTAACAATTAATAATAATGGA 1198
Db      1250 AGTGTAAAGACAGCTGATTAAGTTCAAGTAGCCCTTCAATACGCTTAATAAATA 1309
Qy      1199 ATGTTCAACAAGTTGAGTTT 1218
Db      1310 GTTCGAAATTAATTAATTTT 1329

```

Search completed: August 25, 2003, 03:27:35  
 Job time : 423 secs

THIS PAGE BLANK (USPS)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 25, 2003, 03:15:12 ; Search time 98 Seconds  
(without alignments)  
5882.106 Million cell updates/sec

Title: US-09-901-572A-1  
Perfect score: 106  
Sequence: 1 aaaaacatcgatgttcaat.....taaacgtttcaccagtt 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1280.4	98.0	2144	2	US-08-525-742-9
2	1270	97.2	1387	2	US-08-525-742-3
3	1264	96.8	1305	1	US-08-185-851A-1
4	1086.8	83.2	2014	2	US-08-525-742-7
5	216.6	16.6	2369	2	US-08-525-742-1
6	65.6	5.0	3057	4	US-09-601-198-55
7	55	4.2	10640	4	US-09-417-485D-5
8	53.6	4.1	1887	4	US-09-601-198-39
9	53.2	4.0	8700	2	US-08-645-193B-18
10	51.6	4.0	8700	2	US-08-392-625-16
11	51.6	4.0	8700	2	US-08-466-961A-16
12	51.6	4.0	53332	4	US-09-801-861-3
13	50.8	3.9	1851	4	US-09-601-198-51
14	50.6	3.9	14066	4	US-09-601-198-56
15	50	3.8	2043	4	US-09-601-198-181
16	49.4	3.8	615	3	US-08-998-416-186
17	49.2	3.8	837	3	US-08-998-416-288
18	49.2	3.8	8920	2	US-08-446-855A-1
19	49.2	3.8	8920	3	US-09-150-741-1
20	48.6	3.7	636	3	US-08-998-416-1137
21	48.2	3.7	1500	4	US-09-601-198-36
22	48.2	3.7	19124	2	US-08-487-826B-13
23	48	3.7	19124	2	US-08-487-826B-13
24	47.6	3.6	827	3	US-08-998-416-535
25	46.8	3.6	732	3	US-08-998-416-1036
26	46.6	3.6	2427	4	US-09-601-198-70
27	46.4	3.6	20674	4	US-09-641-638-651

28	46.2	3.5	701	3	US-08-998-416-701	Sequence 701, App
29	46.2	3.5	15016	4	US-09-601-198-60	Sequence 60, Appl
30	46	3.5	998	3	US-09-122-400B-5	Sequence 5, Appl
31	45.8	3.5	1431	3	US-09-316-083-2	Sequence 2, Appl
32	45.8	3.5	1431	4	US-09-933-700-2	Sequence 1, Appl
33	45.8	3.5	1830121	4	US-09-557-884-1	Sequence 1, Appl
34	45.8	3.5	1830121	4	US-09-643-990A-1	Sequence 1, Appl
35	45.6	3.5	3399	4	US-09-601-198-43	Sequence 43, Appl
36	45.6	3.5	6124	4	US-08-213-419B-3	Sequence 3, Appl
37	45.4	3.5	945	4	US-09-601-198-177	Sequence 177, App
38	45.4	3.5	3509	4	US-09-255-829-19	Sequence 19, Appl
39	45.4	3.5	202001	4	US-09-734-674-3	Sequence 3, Appl
40	45	3.4	5340	4	US-09-627-122-21	Sequence 21, Appl
41	44.8	3.4	30549	4	US-09-134-001C-322	Sequence 322, App
42	44.6	3.4	6243	2	US-09-056-075-1	Sequence 1, Appl
43	44.6	3.4	6265	4	US-09-129-112-3	Sequence 3, Appl
44	44.6	3.4	1664976	4	US-08-916-421B-1	Sequence 1, Appl
45	44.4	3.4	658	3	US-08-998-416-595	Sequence 595, App

## ALIGNMENTS

RESULT 1  
US-08-525-742-9  
Sequence 9, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Sasaki, Sakiko  
APPLICANT: Ohnawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigema  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEROPE  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSEE: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Ie-Mhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930





```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-887-0357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 202..1308
US-08-185-851A-1

```

```

Query Match      96.8%; Score 1264; DB 1; Length 1305;
Best Local Similarity 98.3%; Pred. No. 2,8e-255;
Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

```

```

QY 1 AAAAATCATGATTTGTTAACTGATATCTTTGCTTAAACAAACAAATCTTCTAACAA 60
DB 1 AAAAATCATGATTTGTTAACTGATATCTTTGCTTAAACAAACAAATCTTCTAACAA 60
QY 61 AATCTTAATTAATTAAGCGGTTAAATTAATACTAAATTAATTAATTAATTAATTAAT 120
DB 61 AATCTTAATTAATTAAGCGGTTAAATTAATACTAAATTAATTAATTAATTAATTAAT 120
QY 121 AACCAAAATCTCTAGTAAATTAACGCTTATTTATTTTATTTTATTTTATTTTATTTT 180
DB 121 AACCAAAATCTCTAGTAAATTAACGCTTATTTATTTTATTTTATTTTATTTTATTTT 180
QY 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 241 TTAGGTACAAACATCTTTCTTCTAGCATTTGGGATTTCTAGCTGTATGCTTATTA 300
DB 241 TTAGGTACAAACATCTTTCTTCTAGCATTTGGGATTTCTAGCTGTATGCTTATTA 300
QY 301 GATGCAAAACCAATTAATGAGCAAAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 GATGCAAAACCAATTAATGAGCAAAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CTAAATCATGTTAAAGCGATGATTAAGCTTCTAGTACCAAGACTATGCAAGATTGA 420
DB 361 CTAAATCATGTTAAAGCGATGATTAAGCTTCTAGTACCAAGACTATGCAAGATTGA 420
QY 421 AGTTATCATGCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 AGTTATCATGCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GAAACAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
DB 481 GAAACAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
QY 541 AAAAGACTTTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 AAAAGACTTTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 ACTTAGAACAACGCTGCTACTAATCTTGAAGTTTGTATCACTGCTTATTAATCAAT 660
DB 601 ACTTAGAACAACGCTGCTACTAATCTTGAAGTTTGTATCACTGCTTATTAATCAAT 660
QY 661 CGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB 661 CGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
QY 721 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 721 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780

```

```

DB 721 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
QY 781 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
DB 781 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
QY 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 AATGTTCACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 960
DB 901 AATGTTCACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 960
QY 961 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
DB 961 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
QY 1021 AACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
DB 1021 AACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
QY 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
DB 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
QY 1141 TTGGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 TTGGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 GTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 CCAGCAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306
DB 1261 CCAGCAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306

```

```

RESULT 4
US-08-525-742-7
Sequence 7, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Sato, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsewa, Ikuroh
APPLICANT: Funato, Hiroo
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemitsu
APPLICANT: Takahashi, Kiyohito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARMSTRONG, WESTERMAN, HAYTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742

```

FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: McLeand, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2014 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 54..1883  
US-08-525-742-7

Query Match 83.2%; Score 1086.8; DB 2; Length 2014;  
Best Local Similarity 96.7%; Pred. No. 2,8e-218;  
Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

150 TTTATTTTATTTTATTTAGT-CATCTTTAAGATATATATATCTTAATTTCTATGATA 208  
1 TTTATTTTATTTTGGTAAATCTTTTAAATATATATATTTTAAATATTTCTATGATA 60  
209 AGAAGAATCATCTTAAGAATCTATTAGTTGGTAAACAACTCTTCTTATGATG 268  
61 AAAAAAGATCATCTTAAGAATCTATTAGTTGGTAAACAACTCTTCTTATGATG 120  
269 GGATTTCTAGCTGTATGTCTATTAATAAAAAAGATGCAAAACCAATATATGCAAAACC 328  
121 GGATTTCTAGCTGTATGTCTATTAATAAAAAAGATGCAAAACCAATATATGCAAAACC 180  
329 AATTGAAGCAGCGGGAATGAGATTACAGATCTTAATCAATGCTTAAGCGATGATG 388  
181 AATTGAAGCAGCGGGAATGAGATTACAGATCTTAATCAATGCTTAAGCGATGATG 240  
389 CTGACTCAAGACATGATGCAAGATTGAAGCTATTCATTCGCTTATGTAAGCTG 448  
241 CTGACTCAAGACATGATGCAAGATTGAAGCTATTCATTCGCTTATGTAAGCTG 300  
449 AAACAGTTAACAATTAACCTTAATGCAACATTAGAAACAATAAATGCTTAAATTT 508  
301 AAACAGTTAACAATTAACCTTAATGCAACATTAGAAACAATAAATGCTTAAATTT 360  
509 TAGAATGAGCCATCAACCAAGCTTAATGAGTAAACAGACTTTTGAATATGAACCCAA 568  
361 TAGAATGAGCCATCAACCAAGCTTAATGAGTAAACAGACTTTTGAATATGAACCCAA 420  
569 AATTAGTTGAAGCATAAAGACATTAATAAACAATTGGAACAGTCTTAACTTGG 628  
421 AATTAGTTGAAGCATAAAGACATTAATAAACAATTGGAACAGTCTTAACTTGG 480  
629 AAGTTTGTCACTCACTGCTTATATCAAAATTCGCAATATTTAGTGAATCAATCAATA 688  
481 AAGTTTGTCACTCACTGCTTATATCAAAATTCGCAATATTTAGTGAATCAATCAATA 540  
689 AAGCTAGTAGTTTAACTAATAAACAATAAAGATGCACTTAATGAGGGAACGCTTTAGATT 748  
541 AAGCTAGTAGTTTAACTAATAAACAATAAAGATGCACTTAATGAGGGAACGCTTTAGATT 600

749 CTAATGAGATTACTACAGTTAATCGAATATTAATAATAGCTTATCACTATTATGAC 808  
601 CTAATGAGATTACTACAGCTTATAGATATTAATAATAGCTTATCACTATTATGAC 660  
809 AAAAGCTAATGCTGATGATCTTATTAATGTTTATTAATAAAGATTCATAATATG 868  
661 AAAAGCTAATGCTGATGATCTTATTAATGTTTATTAATAAAGATTCATAATATATA 720  
869 AACAAATTTTGTGAGGACCTTTTCAAAACGCTAATGTTCAACCTTCAAACTACGTTTG 928  
721 AACAAATTTTGTGAGGATTTTCAAAACGCTAATGTTCAACCTTCAAACTATGTTTG 780  
929 TTGCTTTAGGCTGATGATCAACCGCTCAATTAATAATGCAAGAGACGTTTGA 988  
781 TTGCTTTAGGCTGATGATCAACCGCTTATTAATAATGCAAGAGACGTTTGA 840  
989 ATGCTGATGAACCTTCAAGTGAATTTCTTCAACAGAAATGATGATCAAGATGTTCTT 1048  
841 ATGCTGATGAACCTTCAAGTGAATTTCTTCAACAGAAATGATGATGATGTTCTT 900  
1049 GGATTTATGTTTATGCTGAAACAAACGAAATGCAATTTAGTTTATGCACTATGCTC 1108  
901 GAATTTATGTTTATGCTGAAACAAACGAAATGCAATTTAGTTTATGCACTATGCTC 960  
1109 CATCACTGCTTATTTATTTTCCCTTATTAAGTTGTTAAGCAGCTGATGCTAATACG 1168  
961 CATCACTGCTTATTTATTTTCCCTTATTAAGTTGTTAAGCAGCTGATGCTAATACG 1020  
1169 TTGATTTACATACAAATTAATTAATGAAATGTTCAACAAGTTGATGCTTCA 1228  
1021 TTGATTTACATACAAATTAATTAATGAAATGTTCAACAAGTTGATGCTTCA 1080  
1229 CTAGTCAATTAATTAATTAATTAATGAAATGTTCAACAAGTTGATGCTTCA 1288  
1081 CTAGTCAATTAATTAATTAATTAATGAAATGTTCAACAAGTTGATGCTTCA 1140  
1289 AATCGTTTATGCTT 1306  
1141 AATCGTTTATGCTT 1158

RESULT 5  
US-08-525-742-1  
Sequence 1, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohsawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshihazu  
APPLICANT: Aoyama, Shigemitsu  
APPLICANT: Takahashi, Kiyohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSER: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30



Qy 624 CTTGAGAGTTTGTCATCACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATA 683  
Db 2170 TCCCTAATAGAGTGAAGAACATTAAGACGAATTTTCAGAAATTAATTTCTAATACAT 2111  
Qy 684 CAATTAAGCTAGTACTTTAATTAATCAAAAACATGATCCACTAAATGGGGAGCGCTTTT 743  
Db 2110 TTAACAAAAAATTAATTAATTAATGAAGATTAATCAATCATCAAGATGAAGAACTAATGA 2051  
Qy 744 AGATTCTAATGAGTACTAATCACTGATTAATCGAAATTTAATAATAGCTTAATCAATCTATA 803  
Db 2050 ACTAATCTTAATCAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1991  
Qy 804 TGAACAAAAGCTAATGCTGATGATTAATCTAATAGTTTAAATTAATTAATTAATTAATTAAT 863  
Db 1990 TAGAAATTAATTTCTTCAAAAAGATTTTAATTTCAAAATTAATTAATTTTGAAGATTCCT 1931  
Qy 864 TAATGAACAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCA 915  
Db 1930 AAATTCATATATATGCTGATTCATTTATATATATCATCATGACCACTGCA 1879

## RESULT 7

US-09-417-485D-5  
; Sequence 5, Application US/09417485D  
; Patent No. 6541202  
; GENERAL INFORMATION:  
; APPLICANT: Long, David M.  
; APPLICANT: Metz, Anneke M.  
; APPLICANT: Love, Ruchelle A.  
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes  
; FILE REFERENCE: 47714-5009-US  
; CURRENT APPLICATION NUMBER: US/09/417,485D  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 10640  
; TYPE: DNA  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (834)..(7385)  
; OTHER INFORMATION: TERT gene  
; NAME/KEY: unsure  
; LOCATION: (1821)..(1837)  
; OTHER INFORMATION: m at position 1821 = a or c; w at position 1837 =  
; OTHER INFORMATION: a or t. Xaa (amino acid) at position 330 = Leu or  
; OTHER INFORMATION: Ile; Xaa at position 335 = Asp or Gly.  
US-09-417-485D-5

Query Match 4.2%; Score 55; DB 4; Length 10640;  
Best Local Similarity 42.7%; Pred. No. 0.0075;  
Matches 336; Conservative 0; Mismatches 450; Indels 1; Gaps 1;

Qy 111 TTTCTTTCACACCAAAATCTCTAGTAATAAACCTTAATTTATTTATTTTATAGTAT 170  
Db 3102 TTTGTCTTTCACCAATTTTGTAGAAATTAAGATTAAGAAAGAAAGAAACCCAAA 3161  
Qy 171 CTTTAAAGATATAATATATATCTAATATCTATGATAGAAAGAAATCATCTTAAGAC 230  
Db 3162 ATGATTAATATTAATTAAGAAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 3221  
Qy 231 TATTAGTTTGTAGTACACATCTTTCTTAGCATTTGGGATTTTACCTGTATGTAT 290  
Db 3222 AAATATATATTAATTAAGAAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 3281  
Qy 231 TACTAAAGATAGTAAACCCCAATATATGCGCAATCCCAATTAAGAGAGCGCGCAATGA 350  
Db 3282 TTTTCCAAAAAATGATCTCTTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 3341  
Qy 351 GTTAACAGATCTAATCAATGCTAAGCGATGACATTAAGCTTCACTACCAAGATATGCCAA 410

Db 3342 AATATATATTTATTAAGAAAGATATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 3401  
Qy 411 GATTGAAGCTAGATTATCATCTGCTTATAGTAAGCTGAAGCACTTAATACATACCTTA 470  
Db 3402 AAAAGTATAGTAAATTAATTAATCAAAATTAAG-GAAATTAACAAAAAGTTTGAACATA 3460  
Qy 471 TGCAACATTAGAACCACTAAAGAAAGCTAAAGAAAGCTAATTAAGTAAATCAGCCATCAACCAAG 530  
Db 3461 TATTAAAAAATTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3520  
Qy 531 TAATAGGATTAAGCACTTTTGTATTAATGAACACCAAAATTTAGTGAAGCATCAAGC 590  
Db 3521 TCATATGAGATGCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 3580  
Qy 591 ACTAAAAACACTTAAGAACACAGCTGCTACTAATCTGAAGCTTTCATCACTACGCTTA 650  
Db 3581 TTTTATTAATTTTGTCTGCAAGAAAGAAAGCAATTAATTAATTAATTAATTAATTAAT 3640  
Qy 651 TAATCAATTTGCAATTAATTTAGTATCTATACATTAAGCTAGTAAATTAATTAATTAAT 710  
Db 3641 TCAAAATTAAGAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3760  
Qy 711 AACATGATTCACATAATGCGGAGCGTTTATGATTTCAATGAGATTAATCAAGTTAA 770  
Db 3761 TATATATTAAGAAAGCAAGTGAATTTGTTAAAGATTAAGATTAAGATTTTAACTTT 3760  
Qy 771 TCGAATATTAATTAATGCTTATCACTATTAATGAACAAAGCTAATGCTATGAT 830  
Db 3761 AATCAAAATTAAGAAAGTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3820  
Qy 831 ATCTAATAGTTTATTAATTAAGAAAGTCAATTAATTAATTAATTAATTAATTAATTAAT 890  
Db 3821 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3880  
Qy 891 TACAAC 897  
Db 3881 TACAAC 3887

## RESULT 8

US-09-601-198-39/c  
; Sequence 39, Application US/09601198  
; Patent No. 6531583  
; GENERAL INFORMATION:  
; APPLICANT: Caselli, Gail H.  
; APPLICANT: Chen, Elison Y.  
; APPLICANT: Glass, Jennifer S.  
; APPLICANT: Heiner, Cheryl R.  
; APPLICANT: Letkowitz, Elliot  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
; FILE REFERENCE: UAB-13452/22  
; CURRENT APPLICATION NUMBER: US/09/601,198  
; CURRENT FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/073,189  
; NUMBER OF SEQ ID NOS: 181  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 1887  
; TYPE: DNA  
; ORGANISM: Ureaplasma urealyticum  
US-09-601-198-39

Query Match 4.1%; Score 53.6; DB 4; Length 1887;  
Best Local Similarity 43.4%; Pred. No. 0.011;  
Matches 349; Conservative 0; Mismatches 449; Indels 6; Gaps 2;

Qy 82 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 141  
Db 1284 TTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1225





APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process For The Preparation  
TITLE OF INVENTION: Of Chemical Compounds  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,625  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-392-625-16

Query Match 4.0%; Score 51.6; DB 2; Length 8700;  
Best Local Similarity 45.2%; Pred. No. 0.037;  
Matches 189; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
QY 512 AATGAGCATACCAAGGATATAGGATTAAGACCTTTGATTAAGAACCCAAAT 571  
DB 4195 ATTGGAATTTATTAAGTGAAGAGAAATTTTATTAATGCGAAGTTAT 4254  
QY 572 TAGTTGAAGCATACCAAGCATTAACCACTTTAGAACCAAGTCTAATCCTTGAAG 631  
DB 4255 ATCGTAGTAATGACATTAAGAAATTAATTAATTAATTAATTAATTAATTA 4314  
QY 632 GTTGTGATCAACGCTGATTAATCAAAATTCGCAATTAATTAATTAATTAATTA 691  
DB 4315 ATGACTATGAATTTTAAAGAAATTTCCCAATCTTAATTAATTAATTAATTA 4374  
QY 692 CTAGTAGTTTAATTAATAACATAGATCCATTAATGGGGGAGCGTTTATAGTTCTA 751  
DB 4375 TTAGTAATTTTAAATCTTAAGAAACATTAATTAATTAATTAATTAATTAATTA 4434  
QY 752 ATGAGATTACTACAGTTATCGGAATTAATTAATTAATTAATTAATTAATTAATTA 811  
DB 4435 GGATTAATGGCAGTTTAAACATGCGTGTATTAATTAATTAATTAATTAATTAATTA 4494  
QY 812 AGACTAATGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 871  
DB 4495 AAGAAATTTGTTTATCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 4554  
QY 872 AAAAGTTTGTAGGAGCTTTTAAACGCAATGTTCAACCTTCAACTAGAGTTTGT 929  
DB 4555 GTTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4612

RESULT 11

US-08-466-961A-16  
Sequence 16, Application US/08466961A  
Patent No. 5843709  
GENERAL INFORMATION:  
APPLICANT: Enlian, Karl-Dieter  
APPLICANT: Gtz, Friedrich  
APPLICANT: Schnell, No. 5843709bert  
APPLICANT: Augustin, Johannes  
APPLICANT: Engelke, Gernar  
APPLICANT: Rosenstein, Ralf  
APPLICANT: Kaletta, Cortina  
APPLICANT: Klein, Cora  
APPLICANT: Wieland, Bernd  
APPLICANT: Kupke, Thomas  
APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process for the Preparation of  
TITLE OF INVENTION: Chemical Compounds  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,961A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/392,625  
FILING DATE: 22-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/784,234  
FILING DATE: 31-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-466-961A-16

Query Match 4.0%; Score 51.6; DB 2; Length 8700;  
Best Local Similarity 45.2%; Pred. No. 0.037;  
Matches 189; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
QY 512 AATGAGCATACCAAGGATATAGGATTAAGACCTTTGATTAAGAACCCAAAT 571  
DB 4195 ATTGGAATTTATTAAGTGAAGAGAAATTTTATTAATTAATTAATTAATTAATTA 4254  
QY 572 TAGTTGAAGCATACCAAGCATTAACCACTTTAGAACCAAGTCTAATCCTTGAAG 631  
DB 4255 ATCGTAGTAATGACATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 4314  
QY 632 GTTGTGATCAACGCTGATTAATCAAAATTCGCAATTAATTAATTAATTAATTAATTA 691

Db 4315 ATGACTATGAAATTTTAAAAAAGAAATTCGCAATCTGATGATTTCTATTATTAATAAA 4374  
Qy 692 CTGTGATTTTAAATTAACCTGATGATCACTAAATGGGGAACGCTTTAGATTTCA 751  
Db 4375 TTAGTATTTTGAAGAAATCTTAAAAAGACATCAAAAAAGCTATATCTTACCTGCTCA 4434  
Qy 752 ATGAGATTTACTACGATTAATCGGATATTAATATGATGATGATCACTATTATTAAGACAA 811  
Db 4435 GGAATATTTGGCAGTTTATTAACATGCGTTGATATGAATTTTGGATTTAAATCTGAAA 4494  
Qy 812 AGACTAATGCTGATGATTAATCAATGATTTTATTTAAAAAGATTCATAAATTAATGAAC 871  
Db 4495 AAGAAAAATTTGTTTATCTATTTTAAATGAATTCAAAACTAAAAATATTTGGGATG 4554  
Qy 872 AAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAAACTACGATTTGT 929  
Db 4555 GTTGTATTAATTAATTAATTAATCAATTAATAAATTTTGAAGAAATTAATCACTTTTGT 4612

## RESULT 12

US-09-801-861-3/c  
Sequence 3, Application US/09801861  
Patent No. 6452154

GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CLO01098  
CURRENT APPLICATION NUMBER: US/09/801,861  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 53332  
TYPE: DNA  
ORGANISM: Human  
US-09-801-861-3

Query Match 4.0%; Score 51.6; DB 4; Length 53332;  
Best Local Similarity 50.7%; Pred. No. 0.052;

Matches 149; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

Qy 24 ATATCTTGGCTTAAAAAACAACAATCTCTTAACAAATCCTTAATATAATAGCCGTTA 83  
Db 31619 AATATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAAT 31560  
Qy 84 AATTAACATAAATAAATAAATAAATGTTTCTTAATCAACAAATCTCTAGTATATAA 143  
Db 31559 AATATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 31500  
Qy 144 CGCTTATTTAT-TTTATTTTATGTCATCTTTAAGATATAATATAATATAATTTCTA 202  
Db 31499 TATTAATATATAATATAATATAATATAATATAATATAATATAATATAATATAAT 31440  
Qy 203 TGAATAGAAAAAGATCACTTAAGACTAATAGTTGTTAGTACATCATCTTTCTTA 262  
Db 31439 TATATAAATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 31380  
Qy 263 GCATTGGATTTCTAGCTGATGCTATTAATCAAAAAAGATGCAACCCAAATA 316  
Db 31379 AATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 31326

## RESULT 13

US-09-601-198-51/c  
Sequence 51, Application US/09601198  
Patent No. 6531583

GENERAL INFORMATION:  
APPLICANT: Caselli, Gail H.  
APPLICANT: Chen, Ellison Y.  
APPLICANT: Chen, Jennifer S.  
APPLICANT: Glass, John I.

APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
TITLE OF INVENTION: UREALYTICUM  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 51  
LENGTH: 1851  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-51

Query Match 3.9%; Score 50.8; DB 4; Length 1851;  
Best Local Similarity 46.4%; Pred. No. 0.041;

Matches 166; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy 533 ATAGGATTAACACGCTTTGATATGATGACACCCAAATTTAGTTGAAGCATACAAAGCAC 592  
Db 794 AAAAAAGTAAACAAATTTCTCAAGCACATATATGTTTGAACACCAATTAATAAGTT 735  
Qy 593 TAAAGCATTGAAACAGCTGCTACTAATCTTGAAGTTGTCATCACTGCTTATA 652  
Db 734 TAAATCAATCATTAATCTTTGATTTTGAATTTTAACTTAAATTAATCAATGATGATCT 675  
Qy 653 ATCAATTTGCAATTAATTTAGTATGATCTATACATTAAGTATGATTAATTAATTAATAA 712  
Db 674 ATATCTTAATGATGATTAATTAATAAATTTCAAAAAGCTTAATTCACAAATTAATGATG 615  
Qy 713 CACTAGATCACTAATGAGGAGACGCTTTAGATTTCTAATGATTAATCACTAATATC 772  
Db 614 TTGATTAATAATTAACAAACCATTTATCAATTTTAAACAAATTAATTAATTAATAGT 555  
Qy 773 GGAATTAATTAATGATTTCACTAATTAATGAACAAAGATTAATGATGATCATTTAT 832  
Db 554 ACTCATTAATAACATTTAAAGAGTTTAATTAATTAATTAATTAATTAATTAATTAAT 495  
Qy 833 CTATAGTTTATTAATAAAGATTCAAATATATGAACAAAGTTTGTAGGACCTT 890  
Db 494 AAAAACTATGCAATTCACAAACATTCAGAAAGCTTTAAAGATTTATGAAGAGATTT 437

## RESULT 14

US-09-601-198-56/c  
Sequence 56, Application US/09601198  
Patent No. 6531583

GENERAL INFORMATION:  
APPLICANT: Caselli, Gail H.  
APPLICANT: Chen, Ellison Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
TITLE OF INVENTION: UREALYTICUM  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 56  
LENGTH: 14066  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-56

Query Match 3.9%; Score 50.6; DB 4; Length 14066;

Best Local Similarity 44.5%; Pred. No. 0.066;  
Matches 200; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

QY 461 ATACCTTAATGCAACATTAGACAACTAATAAAGCGCTAAACCTAATTAGATGAGCA 520  
DB 10561 ATACATCAATTTATCATTTATTAAGAGCGCTGTAAGCAAGATATTTGACAGCATG 10502  
QY 521 TCACCAAGCTAATACGATTAAGACCTTTGATATGAAACACCAATTTAGTTGAAG 580  
DB 10501 CTAATTTATTTATGATTTATTAAGATTTGATGATCAAAATTTGATTAAGATGTTA 10442  
QY 581 CATACAAAGCACTAAACACCTTTAGAACAAAGCTGCTACTACCTTGAAGTTTGCAT 640  
DB 10441 AATTTATTTATCAAAATTTGATTAACATAAAGCACTCAAAATTTGATATGAAAG 10382  
QY 641 CAACGCTTATTAATGCAATTTAGTGAATCTATACATTAAGTATGATTT 700  
DB 10381 TTGTTGCAAAATTAATAATTAATTCATTTGTTGTTTAAAGAAAAACAACATATG 10322  
QY 701 TAATTAATAAACAATAATGCAATTAATGAGGAGCGCTTTTATGATTTATATGATTA 760  
DB 10321 TTATTAATAAAGCTTGAAGCTTTAATTAAGATGAGAGATTTGTTATGATGATTTG 10262  
QY 761 CTACAGTTAATGCAATTTATTAATTAATGATTAATCACTAATTAATGACAAAGACTAATG 820  
DB 10261 ATTACTAGATCCAAATCTAATTTTATGATCACTAATAAACAACAACATAGGCTC 10202  
QY 821 CTGAGCACTAATCTAATGTTTATTAATAAAGATTCATCAAAATATGACAAAGTTTG 880  
DB 10201 TAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10142  
QY 881 TAGGACTTTTACAAAGCTAATGTTCAA 909  
DB 10141 TAGTGATGATTTTCAAGATATCAACA 10113

## RESULT 15

US-09-601-198-181

Sequence 181, Application US/09601198

Patent No. 6531583

GENERAL INFORMATION:

APPLICANT: Caselli, Gail H.

APPLICANT: Chen, Ellison Y.

APPLICANT: Glaes, Jennifer S.

APPLICANT: Glaes, John I.

APPLICANT: Heiner, Cheryl R.

APPLICANT: Lefkowitz, Elliot

TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA

FILE REFERENCE: UAB-13452/22

CURRENT APPLICATION NUMBER: US/09/601,198

PRIOR FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/073,189

NUMBER OF SEQ ID NOS: 181

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 181

LENGTH: 2043

TYPE: DNA

ORGANISM: Ureaplasma urealyticum

US-09-601-198-181

Query Match 3.8%; Score 50; DB 4; Length 2043;  
Best Local Similarity 42.3%; Pred. No. 0.061;

Matches 348; Conservative 0; Mismatches 465; Indels 9; Gaps 1;

QY 442 GAAGCTGAACAGTTAATCACTTAATGCAATTAAGCACTAATAAATGCTAAA 501  
DB 1126 GGATATGAATTAATTAATAATGATCATATGCAATTAACAGAAATTAATAAAT 1185  
QY 502 ACTAATTTAGATCA-----GCCATCAACCAAGCTAATACGATTAACGACTTTT 552  
DB 1186 ACCATTTTAGAACCAAAAGCTTGTGCAACCAAGATGATCATGATTTAAAGTTAAAT 1245

QY 553 GATATGAACCCCAATTTAGTTGAAGCATACAAAGCACTAATAAACCATTGACAA 612  
DB 1246 GATATGAAGAGAAATGAAGATTTTGGACAAATCTGATTAATTAAGAAATCAAGAT 1305  
QY 613 CGTGCTAATCACTTGAAGTTTGTATCAATGCTCTTAATTAATCAAAATTCGATTAATTTA 672  
DB 1306 TCTACAGTCTGTTGAATTAATTAATGATTCATCTGAAGTACTCAAGATGATTAACGAT 1365  
QY 673 GTGATCTAATCAATTAAGCTAGATTTAATACTAAACATGATCCATTAATGGG 732  
DB 1366 GAAATGTGATTAATTTAATTAATAAGCAATTAACCTGTTTAACGATCTTAACAA 1425  
QY 733 GAAGCTTTTATGATCTAATGAGATTAATCACTAATGCAATTAATTAATACCTTA 792  
DB 1426 GAGATCATCTCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1485  
QY 793 TCACATTAATGAACAAAGACTAATGCTGATGATTAATCTAATTAATTAATTAATTA 852  
DB 1486 GATGAATTTAATTAATGATGATGAAGAAAGATTAATTAATTAATTAATTAATTAAT 1545  
QY 853 GTGATTAATTAATGAACAAAGCTTTGATGAGCTTTTACAAACGCTAATGTTCAACT 912  
DB 1546 ATGATGATGATTAATGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1605  
QY 913 TCACATTAAGTTTGTGCTTTTATGCTGATGATTAACCCGCAATTAATTAATTAATGA 972  
DB 1606 CGAATTAAGATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1665  
QY 973 AGAAGACCGTTGGAATGATGAACCTTCAAGTGAATTTCTTGCAACACGAATAGT 1032  
DB 1666 ATTAATACCAAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1725  
QY 1033 ATCAAGATTTTGTGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1092  
DB 1726 TATGAACCTGAAGAAATGTTAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 1785  
QY 1093 TTGACAACTATGCTCCATCACTGTTAATTTATTTCCCTTAATGTTGTTAAAGCA 1152  
DB 1786 TCAATTAACCAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1845  
QY 1153 GCTGATGCTAATAAGCTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1212  
DB 1846 AAGAAATATCTTTTAACGCCAAAGAAATGAATTAATTAATTAATTAATTAATTAATTA 1905  
QY 1213 GAGTTGCCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1254  
DB 1906 GAATTAACACCTGCTGATTTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 1947

Search completed: August 25, 2003, 05:45:37  
Job time : 109 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 03:22:22 ; Search time 1244 Seconds  
(without alignments)  
2360.211 Million cell updates/sec

Title: US-09-901-572A-1

Perfect score: 1306  
Sequence: 1 aaaaacatcgatgttcaat.....taaacgtttatcaggtc 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	78.7	1152	14	US-10-131-591A-1
2	1024.2	78.4	1371	9	US-09-147-052-1
3	1024.2	78.4	3261	9	US-09-147-052-3
4	921.6	70.6	1082	14	US-10-131-591A-24
5	73.4	5.6	4985	14	US-10-094-240-10
6	68.6	5.3	8136	12	US-10-311-455-128
7	68.4	5.2	11836	12	US-10-240-453-114
8	68.4	5.2	11836	14	US-10-239-676-102
9	68.2	5.2	8392	12	US-10-311-455-1463
10	68	5.2	12405	12	US-10-240-453-43
11	68	5.2	12405	14	US-10-239-676-35
12	66.4	5.1	6669	12	US-10-311-455-166
13	66.4	5.1	6669	12	US-10-311-455-1221
14	65.6	5.0	5163	12	US-10-311-455-1797
15	65.4	5.0	8771	12	US-10-311-455-1
16	64.8	5.0	7351	12	US-10-311-455-1

C 17	64.8	5.0	18512	12	US-10-311-455-950	Sequence 950, App
C 18	64.2	4.9	17594	12	US-10-311-455-1939	Sequence 1939, Ap
C 19	63.2	4.8	16766	12	US-10-311-455-2130	Sequence 2130, Ap
C 20	63	4.8	9254	12	US-10-240-453-92	Sequence 92, App1
C 21	63	4.8	9254	14	US-10-239-676-86	Sequence 86, App1
C 22	62.6	4.8	5883	12	US-10-311-455-326	Sequence 326, App1
C 23	62.6	4.8	5883	12	US-10-240-485-26	Sequence 26, App1
C 24	62.2	4.8	5678	12	US-10-311-455-1111	Sequence 1111, Ap
C 25	62.2	4.8	6061	12	US-10-311-455-114	Sequence 114, App
C 26	62.2	4.8	9728	12	US-10-311-455-1876	Sequence 1876, App
C 27	62	4.7	5520	12	US-10-311-455-1492	Sequence 1492, Ap
C 28	61.4	4.7	7001	12	US-10-172-086-60	Sequence 60, App1
C 29	61.2	4.7	5152	12	US-10-204-708-73	Sequence 73, App1
C 30	61.2	4.7	5421	12	US-10-311-455-105	Sequence 105, App
C 31	61	4.7	6831	12	US-10-311-455-1460	Sequence 1460, App
C 32	60.8	4.7	15387	12	US-10-311-455-157	Sequence 157, App
C 33	60.6	4.6	5930	12	US-10-311-455-1368	Sequence 1368, App
C 34	60.4	4.6	5145	12	US-10-311-455-321	Sequence 321, App
C 35	60.4	4.6	5145	12	US-10-240-485-17	Sequence 17, App1
C 36	60.4	4.6	5641	12	US-10-311-455-1369	Sequence 1369, App
C 37	60.4	4.6	11996	12	US-10-240-485-45	Sequence 45, App1
C 38	60.2	4.6	7461	12	US-10-311-455-1758	Sequence 1758, Ap
C 39	60.2	4.6	9515	12	US-10-240-453-182	Sequence 182, App
C 40	60.2	4.6	9515	14	US-10-239-676-160	Sequence 160, App
C 41	60	4.6	6050	12	US-10-311-455-1984	Sequence 1984, Ap
C 42	60	4.6	9510	12	US-10-240-485-115	Sequence 115, App
C 43	60	4.6	9964	12	US-10-311-455-71	Sequence 71, App1
C 44	60	4.6	17967	12	US-10-311-455-988	Sequence 988, App
C 45	59.4	4.5	11049	12	US-10-204-708-23	Sequence 23, App1

ALIGNMENTS

RESULT 1  
US-10-131-591A-1  
; Sequence 1, Application US/10131591A  
; Publication No. US20030059799A1  
; GENERAL INFORMATION:  
; APPLICANT: Nippon Zeon Co., Ltd.,  
; TITLE OR INVENTION: Modified DNA molecules, Recombinants and uses thereof  
; FILE REFERENCE: J209  
; CURRENT APPLICATION NUMBER: US/10/131,591A  
; CURRENT FILING DATE: 2002-08-15  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1152  
; TYPE: DNA  
; ORGANISM: Mycoplasma gallisepticum  
; FEATURE:  
; OTHER INFORMATION: TTM-1 gene (after EcoRI)  
US-10-131-591A-1

Query Match	78.7%	Score 1028;	DB 14;	Length 1152;
Best Local Similarity	100.0%;	Pred. No. 3.8e-170;		
Matches 1028;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	279	CTGTATGCTTATCTAATAAAGATGCAACCCCAATATATGCGCAAAATTAAGAGC	338	
DB	6	CTGTATGCTTATCTAATAAAGATGCAACCCCAATATATGCGCAAAATTAAGAGC	65	
QY	339	AGCGGAAATGAGTTAACAATCAATGCTTAAGCGATGACATTAGGCTACTACA	398	
DB	66	AGCGGAAATGAGTTAACAATCAATGCTTAAGCGATGACATTAGGCTACTACA	125	
QY	399	AGACTATCCCAAGATTAGCTTATATCATCTGCTTATATAGTAAGCTGAACAGTTAA	458	
DB	126	AGACTATCCCAAGATTAGCTTATATCATCTGCTTATATAGTAAGCTGAACAGTTAA	185	
QY	459	CAATTAACCTTAATGCACTTATGAACCACTTAAATGGCTTAAATTAATTAATGATAC	518	
DB	186	CAATTAACCTTAATGCACTTATGAACCACTTAAATGGCTTAAATTAATTAATGATAC	245	

QY 519 CATCAACGAGTAATACGATGATAAAGCACTTTGATTAATGAACACCCAAATTTAGTGA 578  
DB 246 CATCAACGAGTAATACGATGATAAAGCACTTTGATTAATGAACACCCAAATTTAGTGA 305  
QY 579 AGCATCAAAAGCACTAAACCACTTGAACAAGTGCACCTGACCTGAGGTTGTC 638  
DB 306 AGCATCAAAAGCACTAAACCACTTGAACAAGTGCACCTGACCTGAGGTTGTC 365  
QY 639 ATCACTGCTTAATCAAAATTCGCAATATTTAGTGATCTATACAAATTAAGCTAGAG 698  
DB 366 ATCACTGCTTAATCAAAATTCGCAATATTTAGTGATCTATACAAATTAAGCTAGAG 425  
QY 699 TTTAATCACTAAACCACTGATCACTAATGGGGGAAGCTTTTGATTTGATTAAGT 758  
DB 426 TTTAATCACTAAACCACTGATCACTAATGGGGGAAGCTTTTGATTTGATTAAGT 485  
QY 759 TACTACAGTTAATCGGATATTAATATACGTTATCACTATTAATGAACAAAGACTAA 818  
DB 486 TACTACAGTTAATCGGATATTAATATACGTTATCACTATTAATGAACAAAGACTAA 545  
QY 819 TCGTATGATTAATCTAATAGTTTATTAATAAAGTATCAAAATTAATGAACAAAGTT 878  
DB 546 TCGTATGATTAATCTAATAGTTTATTAATAAAGTATCAAAATTAATGAACAAAGTT 605  
QY 879 TGAAGGACTTTTACAAAGCTATGTCACCTTCAACTCAAGTTTGTGCTTTAG 938  
DB 606 TGAAGGACTTTTACAAAGCTATGTCACCTTCAACTCAAGTTTGTGCTTTAG 665  
QY 939 TCGTATGATTAATCGGATATTAATATACGAGAGAGAGCGTTGGATGATGATGA 998  
DB 666 TCGTATGATTAATCGGATATTAATATACGAGAGAGAGCGTTGGATGATGATGA 725  
QY 999 ACCCTCAATGAATTTCTTGCAACACGAAATGATCAAGATGTTCTTGATTTAAG 1058  
DB 726 ACCCTCAATGAATTTCTTGCAACACGAAATGATCAAGATGTTCTTGATTTAAG 785  
QY 1059 TTTAGTGAACAAACAGAGTACCAATTTAGTTTACCACTATGCGTCAACAGG 1118  
DB 786 TTTAGTGAACAAACAGAGTACCAATTTAGTTTACCACTATGCGTCAACAGG 845  
QY 1119 TTAATTAATTTCCCTTAATAGTTGTTAAGCAGCTGATGCTAATTAAGTTGATTA 1178  
DB 846 TTAATTAATTTCCCTTAATAGTTGTTAAGCAGCTGATGCTAATTAAGTTGATTA 905  
QY 1179 ATACAATTAATTAATGAATGTTCAACAACTGATGCTGCACTTCACTAGTGCAA 1238  
DB 906 ATACAATTAATTAATGAATGTTCAACAACTGATGCTGCACTTCACTAGTGCAA 965  
QY 1239 TAACTACAGCTAATCCAACTCCAGAGTGTGATTAAGTTGCTAAATGCTTT 1298  
DB 966 TAACTACAGCTAATCCAACTCCAGAGTGTGATTAAGTTGCTAAATGCTTT 1025  
QY 1299 ATCAGGTT 1306  
DB 1026 ATCAGGTT 1033

## RESULT 2

US-09-147-052-1  
Sequence 1, Application US/09147052  
Patent No. US20010014335A1  
GENERAL INFORMATION:  
APPLICANT: SAITOH, Shuji  
APPLICANT: YANAGIDA, Yoshiaki  
APPLICANT: TSUZUKI, Yoshinari  
TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,  
FILE REFERENCE: 981167  
CURRENT APPLICATION NUMBER: US/09/147, 052  
CURRENT FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: JP 08-103548  
PRIOR FILING DATE: 1996-03-29

PRIOR APPLICATION NUMBER: PCT/JP97/01084  
PRIOR FILING DATE: 1997-03-28  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1371  
TYPE: DNA  
ORGANISM: hybrid  
US-09-147-052-1  
Query Match 78.4%; Score 1024.2; DB 9; Length 1371;  
Best Local Similarity 99.7%; Pred. No. 1.8e-169;  
Matches 1026; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 278 GCTGATGCTAATTAATCAAAAGATGCAACCCAAATTAATGAGCCAAATTTAGAG 337  
DB 188 GCTGATGCTAATTAATCAAAAGATGCAACCCAAATTAATGAGCCAAATTTAGAG 247  
QY 338 CAGGCGAATGAGTTAACAATCAATGCTAAAGGATGACATTAAGTCACTAC 397  
DB 248 CAGGCGAATGAGTTAACAATCAATGCTAAAGGATGACATTAAGTCACTAC 307  
QY 398 AAGACTATGCCAATGTAAGCTATGATCATGCTTATAGTGAAGCTGAACAGTTA 457  
DB 308 AAGACTATGCCAATGTAAGCTATGATCATGCTTATAGTGAAGCTGAACAGTTA 367  
QY 458 ACAATTAACCTTAATGCAACATTAAGCACTAAATGCTAAATGCTAAATTTAGAAATCAG 517  
DB 368 ACAATTAACCTTAATGCAACATTAAGCACTAAATGCTAAATGCTAAATTTAGAAATCAG 427  
QY 518 CCATCAACGAGCTAATAGGATTAACGCTTTGATTAATGAACACCCAAATTTAGT 577  
DB 428 CCATCAACGAGCTAATAGGATTAACGCTTTGATTAATGAACACCCAAATTTAGT 487  
QY 578 AAGCATCAAAAGCACTAAACCACTTGAACAACGCTGACTAATCTTGAAGTTGT 637  
DB 488 AAGCATCAAAAGCACTAAACCACTTGAACAACGCTGACTAATCTTGAAGTTGT 547  
QY 638 CATCACTGCTTAATTAATCAAAATTTGCAATTAATTTGATCTATCAATTAAGCTAGTA 697  
DB 548 CATCACTGCTTAATTAATCAAAATTTGCAATTAATTTGATCTATCAATTAAGCTAGTA 607  
QY 698 GTTAATTAATCAAAAGCTAGTCACTAATGGGGAGCGTTTGTGATTTCAATAGA 757  
DB 608 GTTAATTAATCAAAAGCTAGTCACTAATGGGGAGCGTTTGTGATTTCAATAGA 667  
QY 758 TTACTACAGTTAATCGGATATTAATATACGTTATCACTATTAATGAACAAAGACTA 817  
DB 668 TTACTACAGTTAATCGGATATTAATATACGTTATCACTATTAATGAACAAAGACTA 727  
QY 818 ATGCTATGATTAATCTAATAGTTTATTAATAAAGTATCAAAATTAATGAACAAAGTT 877  
DB 728 ATGCTATGATTAATCTAATAGTTTATTAATAAAGTATCAAAATTAATGAACAAAGTT 787  
QY 878 TTTAGGGAAGCTTTTCAAAAGCTAATGTTCAACCTTCAAGTTTGTGCTTTA 937  
DB 788 TTTAGGGAAGCTTTTCAAAAGCTAATGTTCAACCTTCAAGTTTGTGCTTTA 847  
QY 938 GTGCTATGTAACACCCGTCATTAATTAATATGCAAGAGAGCGTTTGAATGCTAGT 997  
DB 848 GTGCTATGTAACACCCGTCATTAATTAATATGCAAGAGAGCGTTTGAATGCTAGT 907  
QY 998 AACCTCAATTAATTTCTTGCAACACGAAATGATCAAGATGTTCTTGATTTATA 1057  
DB 908 AACCTCAATTAATTTCTTGCAACACGAAATGATCAAGATGTTCTTGATTTATA 967  
QY 1058 GTTAGTGAACAAACAGAGTACCAATTTGTTTGAACATTAATGCTGATCACTG 1117  
DB 968 GTTAGTGAACAAACAGAGTACCAATTTGTTTGAACATTAATGCTGATCACTG 1027  
QY 1118 GTTAATTAATTTCCCTTAATAGTTGTTAAGCAGCTGATGCTAATTAAGCTGATTA 1177  
DB 1028 GTTAATTAATTTCCCTTAATAGTTGTTAAGCAGCTGATGCTAATTAAGCTGATTA 1087

Query 1178 AATACAAATTAATATGAAATGTTCAACAGTTGACCTTCAACTAGTGCAG 1237  
Db 1088 AATACAAATTAATATGAAATGTTCAACAGTTGACCTTCAACTAGTGCAG 1147  
Query 1238 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATCGTTT 1297  
Db 1148 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATCGTTT 1207  
Query 1298 TATCAGGTT 1306  
Db 1208 TATCAGGTT 1216

RESULT 3  
US-09-147-052-3  
Sequence 3, Application US/09147052  
Patent No. US20010014335A1  
GENERAL INFORMATION:  
APPLICANT: SAITOH, Shuji  
APPLICANT: TSUZAKI, Yoshinari  
APPLICANT: YAMAGUCHI, No. US20010014335A1  
TITLE OF INVENTION: NOVEL FIRED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,  
TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE  
FILE REFERENCE: 981167  
CURRENT APPLICATION NUMBER: US/09/147, 052  
CURRENT FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: JP 08-103548  
PRIOR FILING DATE: 1996-03-29  
PRIOR APPLICATION NUMBER: PCT/JP97/01084  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 3261  
TYPE: DNA  
ORGANISM: hybrid  
US-09-147-052-3

Query Match 78.4%; Score 1024.2; DB 9; Length 3261;  
Best Local Similarity 99.7%; Pred. No. 2,5e-169;  
Matches 1026; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query 278 GCTGATGCTATCTATCTAAAAAGATGCAAAACCAATTAATGCGCAACCAATTAAG 337  
Db 2078 GCTGATGCTATCTATCTAAAAAGATGCAAAACCAATTAATGCGCAACCAATTAAG 2137  
Query 338 CAGCGCGAATGAGTTAAAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCATAC 397  
Db 2138 CAGCGCGAATGAGTTAAAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCATAC 2197  
Query 398 AAGACTATGCCAAGTTGAAGTATCTATCTGCTTATATGTAAGTGAAGTGA 457  
Db 2198 AAGACTATGCCAAGTTGAAGTATCTATCTGCTTATATGTAAGTGAAGTGA 2257  
Query 458 ACAATTAACCTTAATGCAATTAAGCAACTAAATGCTAAATCTAATTAGATGAG 517  
Db 2258 ACAATTAACCTTAATGCAATTAAGCAACTAAATGCTAAATCTAATTAGATGAG 2317  
Query 518 CCATCAACCAAGCTTAATGCGATTAAGCACTTTGATTAATGAACACCAATTTAGTTG 577  
Db 2318 CCATCAACCAAGCTTAATGCGATTAAGCACTTTGATTAATGAACACCAATTTAGTTG 2377  
Query 578 AAGCATTAAGGCTTAATTAAGCACTTTGATTAATGAACACCAATTTAGTTG 637  
Db 2378 AAGCATTAAGGCTTAATTAAGCACTTTGATTAATGAACACCAATTTAGTTG 2437  
Query 638 CATCAACTGCTTATATCAATTCGCAATTAATTTGATGATCTATACATTAAGCTAGTA 697  
Db 2438 CATCAACTGCTTATATCAATTCGCAATTAATTTGATGATCTATACATTAAGCTAGTA 2497  
Query 698 GTTATTAATCACTAAACCTAATGCTTAATTTGAGGAGACGCTTTTATGATTTAATGAGA 757

Db 2498 GTTATTAATCACTAAACCTAATGCTTAATTTGAGGAGACGCTTTTATGATTTAATGAGA 2557  
Query 758 TTACTACAGTTATCGGAATTAATATATGCTATATCAACTTAATGAACAAAGCTA 817  
Db 2558 TTACTACAGTTATCGGAATTAATATATGCTATATCAACTTAATGAACAAAGCTA 2617  
Query 818 ATGCTGATGCTATCTATATGTTTATTTAAAAAGATTCAAATTAATGAACAAAGTT 877  
Db 2618 ATGCTGATGCTATCTATATGTTTATTTAAAAAGATTCAAATTAATGAACAAAGTT 2677  
Query 878 TTGATGAGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTGTGCTTTA 937  
Db 2678 TTGATGAGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTGTGCTTTA 2737  
Query 938 GTGCTGATGTTACACCGTCAATTAATATATGCAAGAGGACGTTTGAATGATG 997  
Db 2738 GTGCTGATGTTACACCGTCAATTAATATATGCAAGAGGACGTTTGAATGATG 2797  
Query 998 AACCTCAAGTGAATTTCTTGCAAAACGATATATCAAGATGTTCTTGATTTATA 1057  
Db 2798 AACCTCAAGTGAATTTCTTGCAAAACGATATATCAAGATGTTCTTGATTTATA 2857  
Query 1058 GTTATGCTGGAACAAACGAGTACCAATTTAGTTTACCACTATGCTTCATCACTG 1117  
Db 2858 GTTATGCTGGAACAAACGAGTACCAATTTAGTTTACCACTATGCTTCATCACTG 2917  
Query 1118 GTTATTAATATTCCTTATATGTTGTTAAAGCGTGAATGCTAATACGTTGATTAAC 1177  
Db 2918 GTTATTAATATTCCTTATATGTTGTTAAAGCGTGAATGCTAATACGTTGATTAAC 2977  
Query 1178 AATACAAATTAATATGAAATGTTCAACAGTTGATTTGCCATTCACATGAGCA 1237  
Db 2978 AATACAAATTAATATGAAATGTTCAACAGTTGATTTGCCATTCACATGAGCA 3037  
Query 1238 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTGCTAAATTCGTTT 1297  
Db 3038 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTGCTAAATTCGTTT 3097  
Query 1298 TATCAGGTT 1306  
Db 3098 TATCAGGTT 3106

RESULT 4  
US-10-131-591A-24  
Sequence 24, Application US/10131591A  
Publication No. US20030059799A1  
GENERAL INFORMATION:  
APPLICANT: Nippon Zeon Co., Ltd.  
TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
FILE REFERENCE: 4209  
CURRENT APPLICATION NUMBER: US/10/131, 591A  
CURRENT FILING DATE: 2002-08-15  
NUMBER OF SEQ ID NOS: 79  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 24  
LENGTH: 1082  
TYPE: DNA  
ORGANISM: Mycoplasma gallisepticum  
FEATURE:  
OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of  
OTHER INFORMATION: pNZ40K-S  
US-10-131-591A-24

Query Match 70.6%; Score 921.6; DB 14; Length 1082;  
Best Local Similarity 98.0%; Pred. No. 1.3e-151;  
Matches 933; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Query 355 ACAATTAATCAATGCTTAAGGATGACATTTGCTTCATCAAGACTAAGCAAT 414  
Db 1 ACAATTAATCAATGCTTAAGGATGACATTTGCTTCATCAAGACTAAGCAAT 60  
Query 415 GAACTAGTTATCATCTGCTTATATGTAAGCTGAAACAGTTAACAATTAACCTTAATGCA 474





Oy 1010 GAATT 1014  
Db 2265 AGCTT 2261

## RESULT 6

US-10-311-455-528/c  
; Sequence 528, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 528  
; LENGTH: 8136  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-528

Query Match 5.3%; Score 68.6; DB 12; Length 8136;  
Best Local Similarity 44.1%; Pred. No. 0.016;  
Matches 375; Conservative 0; Mismatches 471; Indels 4; Gaps 2;

Oy 2 AAACATCAGATGTTGTAATCTGATCTTGGCTTAAAAAACAACAAATCTTCACAAA 61  
Db 1591 AAATACATATTTTATTTATCTATCTACATATATATTAACCCCAAAATACCTTAACAT 1532  
Oy 62 ATCTAAATTAATAAGCGGTAAATTAATTAATAATAATAATAATAATAATAATAATA 121  
Db 1531 TTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1472  
Oy 122 ACCAAATTTCTCTAGTATTAACGCTTATTTATTTTATTTAGTCTTTTAAAGTA 181  
Db 1471 AAAAATTAATTAACATATTAATAATAATAATAATAATAATAATAATAATAATAATA 1412  
Oy 182 TAAATATATCTTAATATCTATGAATAGAAAGATCATCTTAAGATTTAGTTGT 241  
Db 1411 TATACAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1352  
Oy 242 TAGGTACACATCTTTCTTGTAGCATGGGATTTCTAGCTGATGCTATTAATAATAAAG 301  
Db 1351 TAAATCAAAAAAATAAGCTATTAATAATAATAATAATAATAATAATAATAATAATAATA 1292  
Oy 302 ATGCAAAACCAATATATGGCCAAACCAATTAAGAGCAGCGGATGAGATTAACAGATC 361  
Db 1291 ACAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1232  
Oy 362 TAATCATGCTTAAGCATGATGATGCTTCAACAGATGATGCTTAAGATGAGCTA 421  
Db 1231 TTTAAAAACCTTATTTATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 1172  
Oy 422 GTTATCATCTGCTTATAGTGAAGCTGAACAGTTACATTAACCTTAATGCAACATTAG 481  
Db 1171 CAAAACAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1112  
Oy 482 AAGCACTAAATATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 541  
Db 1111 TAAACTTCAAAAAATTTAAA-TAAAAAAACTAGATTAACAAATTAATTAATTAATTAATTA 1053

Oy 542 AACGACTTTGATATGAAACACCAAAATTTAGTGAAGCATACAAAGCACTAAAAACA 601  
Db 1052 AAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 993  
Oy 602 CTTTGAACAACGCTCT--ACTTACCTTGAGGTTGTCTATCACTGCTTATATCAAA 658  
Db 992 TTTCTCAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 933  
Oy 659 TTGCAATTAATTTAGTGATCTATACAAATAAGCTAGATTTATATTAATACTAAACACTAG 718  
Db 932 TTAATAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 873  
Oy 719 ATCCATTAATGCGGGAACGCTTTTGAATCTTAATGATTAATCACTTAATGAGATA 778  
Db 872 TTAACCTTAACATTAACAGCCCTATTAATACTTTTAAATAATCCTAATAATAATCAATAA 813  
Oy 779 TTATATATGCTTATCACTTATTAATGAACAAAGACTAATGCTGATGATTAATTAATA 838  
Db 812 AAATATATTTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 753  
Oy 839 GTTTTATTA 848  
Db 752 AATTAATTA 743

## RESULT 7

US-10-240-453-114/c  
; Sequence 114, Application US/10240453  
; Publication No. US20030148326A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated with DNA Transcription  
; FILE REFERENCE: 5013.1009  
; CURRENT APPLICATION NUMBER: US/10/240,453  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03973  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 350  
; SEQ ID NO 114  
; LENGTH: 11836  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (7603)  
US-10-240-453-114

Query Match 5.2%; Score 68.4; DB 12; Length 11836;  
Best Local Similarity 44.3%; Pred. No. 0.02;  
Matches 371; Conservative 0; Mismatches 461; Indels 6; Gaps 2;

Oy 37 AAAAAACAACAATCTTCTAACAATAATCTTAATAATAAGCCCTTAATTAACCTAAAAA 96  
Db 4357 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4298  
Oy 97 ATTAATAAATATGTTTTCTTATCAACCAAAATCTCTAGTAATAACGCTTATTTATTT 156  
Db 4297 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4238

157 TTATTTTGTGATCTTTTAAATATATATATCTTAATCTTAATGATAGAAAAGA 216  
4237 TAAATTTTAT 4178  
217 ATCATCTTAAAGATATATATATATATATATATATATATATATATATATAT 276  
4177 ATAAAAAACAACATATATATATATATATATATATATATATATATATATAT 4120  
277 AGCTGATGCTATTTCTTAAATATATATATATATATATATATATATATAT 336  
4119 AT 4060  
337 GCAGCGGATGAGATTAACAGATCTAATCAATGCTAAAGCATGATAGTCACTA 396  
4059 ACCATCAACAT 4000  
397 CAAGATATGCAAGATGAGATGATATATATATATATATATATATATATAT 456  
3999 TCAAT 3944  
457 AACAT 516  
3943 AAAAAAAT 3884  
517 GCATCAACCAAGCTATATATATATATATATATATATATATATATATATAT 576  
3883 ACTTGAT 3824  
577 GAAGATCAAAAGCACTTAAATATATATATATATATATATATATATATAT 636  
3823 CTAT 3764  
637 TCATCAACCTGATATATATATATATATATATATATATATATATATATAT 696  
3763 AATTCCTAT 3704  
697 AGTTTAT 756  
3703 AAAAAAAT 3644  
757 ATTACTACGTTAT 816  
3643 ATTAT 3584  
817 AATGCTGATGAT 874  
3583 AAAAACTAT 3526

RESULT 8  
US-10-239-676-102/c  
Sequence 102, Application US/10239676  
Publication No. US20030082609A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PISENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
FILE REFERENCE: 5013.1003  
CURRENT APPLICATION NUMBER: US/10/239,676  
CURRENT FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: PCT/EP01/03968  
DE 10019058.8  
DE 10019173.8  
DE 10032529.7  
DE 10043826.1  
PRIOR FILING DATE: 2001-04-06  
2000-04-06  
2000-04-07  
2000-06-30  
2000-09-01  
NUMBER OF SEQ ID NOS: 228

SEQ ID NO 102  
LENGTH: 11836  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
PEPTIDE:  
NAME/KEY: unsure  
LOCATION: (7603)  
US-10-239-676-102

Query Match  
Best Local Similarity 44.3%; Pred. No. 0.02; DB 14; Length 11836;  
Matches 371; Conservative 0; Mismatches 461; Indels 6; Gaps 2;

37 AAAAAACAATAATCTTCTTAAACAAATCTTAAATATATATATATATATATATAT 96  
4357 ATAAAT 4298  
97 ATTAATAAAT 156  
4297 ATAAAT 4238  
157 TTATTTTGTGATCTTTTAAATATATATATCTTAATCTTAATGATAGAAAAGA 216  
4237 TAAATTTTAT 4178  
217 ATCATCTTAAAGAT 276  
4177 ATAAAAAACAACAT 4120  
277 AGCTGATGCTATTTCTTAAATATATATATATATATATATATATATATATAT 336  
4119 AT 4060  
337 GCAGCGGATGAGATTAACAGATCTAATCAATGCTAAAGCATGATAGTCACTA 396  
4059 ACCATCAACAT 4000  
397 CAAGATATGCAAGATGAGATGATATATATATATATATATATATATATATAT 456  
3999 TCAAT 3944  
457 AACAT 516  
3943 AAAAAAAT 3884  
517 GCATCAACCAAGCTATATATATATATATATATATATATATATATATATATAT 576  
3883 ACTTGAT 3824  
577 GAAGATCAAAAGCACTTAAATATATATATATATATATATATATATATATAT 636  
3823 CTAT 3764  
637 TCATCAACCTGAT 696  
3763 AATTCCTAT 3704  
697 AGTTTAT 756  
3703 AAAAAAAT 3644  
757 ATTACTACGTTAT 816  
3643 ATTAT 3584  
817 AATGCTGATGAT 874  
3583 AAAAACTAT 3526

RESULT 9

US-10-311-455-1463/c  
 ; Sequence 1463, Application US/10311455  
 ; Publication No. US20030143606a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIERENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
 ; FILE REFERENCE: 5013.1014  
 ; CURRENT APPLICATION NUMBER: US/10/311,455  
 ; PRIOR FILING DATE: 2002-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
 ; PRIOR FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: DE 10032529.7  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: DE 10043826.1  
 ; PRIOR FILING DATE: 2000-09-01  
 ; NUMBER OF SEQ ID NOS: 2424  
 ; SEQ ID NO 1463  
 ; LENGTH: 8392  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-311-455-1463

Query Match 5.2%; Score 68.2; DB 12; Length 8392;

Best Local Similarity 45.3%; Pred. No. 0.019; Mismatches 443; Indels 16; Gaps 3;

Matches 380; Conservative 0; Mismatches 443; Indels 16; Gaps 3;

36 AAAAAAAAAAAGTCTTCTTACCAAAATCTTAAATTAATAGCCCTTAATTAATTAATAA 95  
 6745 AAAAAAAAAAAGTCTTCTTACCAAAATCTTAAATTAATAGCCCTTAATTAATTAATAA 6686  
 96 AATTAATAAATGCTTTCTTATCAACCAAAATCTTCTAGTAATAAGCTTATTATT 155  
 6685 AAAAAAAAAAAGTCTTCTTATCAACCAAAATCTTCTAGTAATAAGCTTATTATT 6626  
 156 TTTATTTTATGATCTTCTTAAAGATATAATATCTTAATTAATCTTGAATTAAGAAAG 215  
 6625 AAAAAAAAAAAGTCTTCTTAAAGATATAATATCTTAATTAATCTTGAATTAAGAAAG 6566  
 216 AATCATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 275  
 6565 AATCAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6506  
 276 TAGCTGATGTTCTTATTAATAAAGATGCAACCAAAATTAATGCGCAACCAATTAGA 335  
 6505 AATCAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6446  
 336 AGCAGCGGATGAGTTTAAAGATCTTAACTTAATCTTAAAGATGAGATTAAGCTTACT 395  
 6445 AATCTTCTTAAATTAATTAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6386  
 396 ACAAGATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 455  
 6385 CCAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6330  
 456 TAAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 515  
 6329 TATTCATCAAC- AATTAATTAATTAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6272  
 516 AGCAGTCAACCAAGTATTAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 575  
 6271 AATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6222  
 576 TGAAGATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 635  
 6221 TTAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 6162  
 636 GTATCAACCTGTTTAAATTAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 695

Db 6161 ACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6102  
 Oy 696 TAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 755  
 Db 6101 AAAAAAAAAAAGTCTTCTTATCAACCAAAATCTTCTAGTAATAAGCTTATTATT 6042  
 Oy 756 GATTACTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 815  
 Db 6041 AATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 5982  
 Oy 816 TAAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 874  
 Db 5981 TCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5923

RESULT 10  
 US-10-240-453-43/c  
 ; Sequence 43, Application US/10240453  
 ; Publication No. US20030148326A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIERENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
 ; TITLE OF INVENTION: Transcription  
 ; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
 ; FILE REFERENCE: 5013.1009  
 ; CURRENT APPLICATION NUMBER: US/10/240,453  
 ; PRIOR FILING DATE: 2002-10-02  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03973  
 ; PRIOR FILING DATE: 2001-04-06  
 ; PRIOR APPLICATION NUMBER: DE 10019058.8  
 ; PRIOR FILING DATE: 2000-04-06  
 ; PRIOR APPLICATION NUMBER: DE 10019173.8  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: DE 10032529.7  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: DE 10043826.1  
 ; PRIOR FILING DATE: 2000-09-01  
 ; NUMBER OF SEQ ID NOS: 350  
 ; SEQ ID NO 43  
 ; LENGTH: 12405  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; NAME/KEY: unsure  
 ; LOCATION: (7895)  
 US-10-240-453-43

Query Match 5.2%; Score 68; DB 12; Length 12405;  
 Best Local Similarity 43.3%; Pred. No. 0.024;  
 Matches 365; Conservative 0; Mismatches 475; Indels 2; Gaps 1;

35 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 94  
 9308 TTAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 9249  
 95 AATTAATAAATGTTTCTTATCAACCAAAATCTTCTAGTAATAAGCTTATTATTAT 154  
 9248 AATTAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 9189  
 155 TTTTATTTTATGATCTTCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 212  
 9188 TTTTAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9129  
 213 AAGATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 272  
 9128 CAAATCTTAAAGATATAATGTTGTTAGTACCAATCTTCTTACATTTGGGATTTTC 9069  
 273 TTCTAGCTGATGTTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 332

Db 9068 TCATATCTTCAACGATTAACATTTATACCAACAAATTAATCCACTAATTTTACAT 9009  
Qy 333 AGAGACGCGCGGATGAGTTAAACAGATCTAATCAATGCTAAAGCAGTACATTTAGCTTC 392  
Db 9008 AAAAAATTAACATTAATAAAAAATTAATAAACCACTTAATAAAAAATTTAAATCCAAA 8949  
Qy 393 ACTACAGAGCTATGCGAGTTGAGCTGTTTATATCTGCTTAATAGGAAGCTGAAC 452  
Db 8948 AATTAATAAACCACTTAACCTTAATCAAAAAATCAAAAAATTAATAATCTTAAAA 8889  
Qy 453 AGTTAACATTAACCTTAATGCAACATTAAGACAACTAAAAATGCTTAATAATTAGA 512  
Db 8888 TCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8829  
Qy 513 ATGAGCATCAACCAAGCTTAATGAGATTAAGCACTTTGATTAATGAACCCCAATT 572  
Db 8828 AATTAATACTTAATCAATTAATTTTACCTTAAAAAATCAATAATTAATCAAAACTA 8769  
Qy 573 AGTTAGAGCATCAACCAAGCTTAATAACCACTTTAGACAAAGCTTACTTAACCTGAAG 632  
Db 8768 TTTTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8709  
Qy 633 TTTGTCATCACTCTTAATTAATCAAAATTCGCAATATTTAGTGAATCTATACATAAAGC 692  
Db 8708 AAAACTAAATTAACCAACTCTTAATAAATAAATAAATAAATAAATAAATAAATAA 8649  
Qy 693 TGTAGCTTAATTAATCAAAACCTTAATCACTTAATGCGGAGAGCTTTAGATTCTAA 752  
Db 8648 AATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8589  
Qy 753 TGAGATTAATCAAGTTATGCGAATTAATTAATTAATTAATTAATTAATTAATTAAT 812  
Db 8588 AACTATTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8529  
Qy 813 GACTAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 872  
Db 8528 AACATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8469  
Qy 873 AA 874  
Db 8468 TA 8467

RESULT 11  
US-10-239-676-35/c  
; Sequence 35, Application US/10239676  
; Publication No. US20030082609A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; PRIOR FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019056.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 35  
; LENGTH: 12405  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (7895)  
US-10-239-676-35  
Query Match  
Best Local Similarity 43.3%; Pred. NO. 0.024;  
Matches 365; Conservative 0; Mismatches 475; Indels 2; Gaps 1;  
5.2%; Score 68; DB 14; Length 12405;  
Qy 35 TAAAAAACACAAAATCTTTACAAAATCCTTAATTAATTAATTAATTAATTAATTAATTA 94  
Db 9308 TTAACATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9249  
Qy 95 AATTAATAAATGAGTTTCTTAATCAACAAAATCTTAATTAATTAATTAATTAATTAATTA 154  
Db 9248 AATTAATCTTAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9189  
Qy 155 TTTTATTTTATGATCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 212  
Db 9188 TTTAAACACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9129  
Qy 213 AAGATCATCTTAAGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 272  
Db 9128 CAATAACACTTAACACGAAATATATATCTTAATAAATAAATAAATAAATAAATAAATAA 9069  
Qy 273 TTTAGCTGATGCTTTTACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 332  
Db 9068 TCAATTAATCTTCAACGAATTAACCAATTTATACCAATAAATAAATAAATAAATAAATA 9009  
Qy 333 AGAGACGCGGATGAGTTAAACAGATCTTAATCAATGCTTAAGCAATTAAGCTTC 392  
Db 9008 AAAAAATTAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8949  
Qy 393 ACTACAGAGCTATGCGAGTTGAGCTGTTTATATCTGCTTAATAGGAAGCTGAAC 452  
Db 8948 AATTAATAAACCACTTAACCTTAATCAAAAAATCAAAAAATTAATAATCTTAAAA 8889  
Qy 453 AGTTAACATTAACCTTAATGCAACATTAAGACAACTAAAAATGCTTAATAATTAGA 512  
Db 8888 TCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8829  
Qy 513 ATGAGCATCAACCAAGCTTAATGCGAATTAATTAATTAATTAATTAATTAATTAATTA 572  
Db 8828 AATTAATACTTAATCAATTAATTTTACCTTAAAAAATCAATAATTAATCAAAACTA 8769  
Qy 573 AGTTAGAGCATCAACCAAGCTTAATAACCACTTTAGACAAAGCTTACTTAACCTGAAG 632  
Db 8768 TTTTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8709  
Qy 633 TTTGTCATCACTCTTAATTAATCAAAATTCGCAATATTTAGTGAATCTATACATAAAGC 692  
Db 8708 AAAACTAAATTAACCAACTCTTAATAAATAAATAAATAAATAAATAAATAAATAA 8649  
Qy 693 TGTAGCTTAATTAATCAAAACCTTAATCACTTAATGCGGAGAGCTTTAGATTCTAA 752  
Db 8648 AATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8589  
Qy 753 TGAGATTAATCAAGTTATGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 812  
Db 8588 AACTATTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8529  
Qy 813 GACTAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 872  
Db 8528 AACATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8469  
Qy 873 AA 874  
Db 8468 TA 8467

RESULT 12  
US-10-204-708-6/c  
; Sequence 6, Application US/10204708

```
Publication No. US20030141852A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPERBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
PRIORITY FILING DATE: 2003-05-06
PRIORITY APPLICATION NUMBER: PCT/EP01/03971
PRIORITY FILING DATE: 2001-04-06
PRIORITY APPLICATION NUMBER: DE 10019058.8
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: DE 10019173.8
PRIORITY FILING DATE: 2000-04-07
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 6
LENGTH: 6669
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-6

Query Match      5.1%; Score 66.4; DB 12; Length 6669;
Best Local Similarity 44.2%; Pred. No. 0.037;
Matches 363; Conservative 0; Mismatches 456; Indels 3; Gaps 2;

Qy 55 TAACAAAATCCTAAATTAATAAGCGTTAAATTAATAATAATAATAATAATGTTT 114
Db 4745 TATCTACACATTAATAAACAATTAATAATAATAATAATAATAATAATAATAAT 4666
Qy 115 CTATCAACCAAAATTCCTAGTATTAACGCTTATTTATTTTATTTTATGTCATCTT 174
Db 4685 CAATTCACAAACACCTTAATAATAATAATAATAATAATAATAATAATAATAAC 4626
Qy 175 TAAGATTAATAATATCTTAATATTTCTATGAATGAAGAAGATCATCTTAAGACT 234
Db 4625 AAACCTTAATATATCTTAACCCACACACATTTTATTAATCCCAACAAACATTA 4566
Qy 235 AGTTGTAGGTACACATCTCTTCTTAGCATGGGATTTCTAGTGTATGTCTATCT 294
Db 4565 AACTACTTAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATA 4506
Qy 295 AAAAAAGATGCAAAACCAATTAATGCGCAAAACCAATTAAGAAGCGCGAATGAGTTA 354
Db 4505 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4446
Qy 355 ACAGATCTAATCAATGCTAAAGCGATGACATTAAGCTTACATCAAGACTATGCCA 414
Db 4445 AAAAAATTAACAATAAACAATAAATATCTTAATTAATAATAATAATAATAATA 4386
Qy 415 GAAGTATGTTATCTCTGCTTATAGTAGAGTGAACAGTTAACAATAAACCCTTA 474
Db 4385 AAATTCACAAATAATTTTTTATTAATATATTTTAAATAATTAATAATAATAATA 4326
Qy 475 ACATTAGAACAATTAATGCTTAAATCTAATTAAGATCAAGCATCAACCAAGCTA 534
Db 4325 TAAACATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4266
Qy 535 ACAGATTAAGACACTTTTGATTAATGAACACCAATTTAGTTGAAGCATACAAAGCA 594
Db 4265 ACATTAACACACCAAAATTTACTTAATAATAATAATAATAATAATAATAATAA 4206
Qy 595 AAAACCACTTTAGAACACAGTGTCTAATCCTGAAGTTGTCTCACTGCTTATAT 654
Db 4205 AAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4146
```

```
Qy 655 CAATTCGCATTAATTAATGATCTATA--CAATAAGCTAGTATTAATACTAA 712
Db 4145 TCATATAATAATAATAATAACGATATATATCTATTAATAATAATAATAATAA 4086
Qy 713 CACTAGATCCACTAAATGGGGAAACGCTTTAGTTCTAATGAGATTACTAGCTAT 772
Db 4085 AATTAATAACATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4026
Qy 773 GGAATTAATAATAATACGTTATCACTATTAATGAACAAAGACTAATGCGATGAT 832
Db 4025 TTTTAATAATAATAATAATAATAATATCTCACTTCAACATTAATAATAATA 3967
Qy 833 CTAAATGTTTATTAATAATAAAGTCAATTAATAATAATAATAATAATAATA 874
Db 3966 CAATAATAATTAATAACACATAATAATAATAATAATAATAATAATAATA 3925

RESULT 13
US-10-311-455-166/c
Sequence 166, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPERBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
PRIORITY FILING DATE: 2002-12-16
PRIORITY APPLICATION NUMBER: PCT/EP01/07537
PRIORITY FILING DATE: 2001-07-02
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 166
LENGTH: 6669
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-166

Query Match      5.1%; Score 66.4; DB 12; Length 6669;
Best Local Similarity 44.2%; Pred. No. 0.037;
Matches 363; Conservative 0; Mismatches 456; Indels 3; Gaps 2;

Qy 55 TAACAAAATCCTAAATTAATAAGCGTTAAATTAATAATAATAATAATAATGTTT 114
Db 4745 TATCTACACATTAATAAACAATTAATAATAATAATAATAATAATAATAATAAT 4666
Qy 115 CTATCAACCAAAATTCCTAGTATTAACGCTTATTTATTTTATTTTATGTCATCTT 174
Db 4685 CAATTCACAAACACCTTAATAATAATAATAATAATAATAATAATAATAATAAC 4626
Qy 175 TAAGATTAATAATATCTTAATATTTCTATGAATGAAGAAGATCATCTTAAGACT 234
Db 4625 AAACCTTAATATATCTTAACCCACACACATTTTATTAATCCCAACAAACATTA 4566
Qy 235 AGTTGTAGGTACACATCTCTTCTTAGCATGGGATTTCTAGTGTATGTCTATCT 294
Db 4565 AACTACTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4506
Qy 295 AAAAAAGATGCAAAACCAATTAATGCGCAAAACCAATTAAGAAGCGCGAATGAGTTA 354
Db 4505 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4446
Qy 355 ACAGATCTAATCAATGCTAAAGCGATGACATTAAGCTTACATCAAGACTATGCCA 414
Db 4445 AAAAAATTAACAATAAACAATAAATATCTTAATTAATAATAATAATAATAATA 4386
```

Qy	415	GAAGCTGTTATATCATCGCTTATAGTGAAGCTGAACAGTTAAACAATTAACCTTAATGA	474
Db	4385	AAATTCAAAAAAATTTTTTTTATATATATATTTTAAATTTTAAAAAACTTAATTATC	4326
Qy	475	ACATTAGAACACTAAAAATGGCTAAAACTAATTTAGATCAGCCATCAACCAAGCTPAT	534
Db	4325	TAAACACTAAAAAAACAAAACTTAACTCTCTATATTTCAACAATTAACAAATTTTAAA	4266
Qy	535	AACGATTAACAGACTTTTGATATATGACACCCCAATTTAGTTGAAGCATACAAAGCACTA	594
Db	4265	ACATATCTAAACACCCAAAAATTTACTTAAAAAAACATTATATAAAAAACCAATPAACAA	4206
Qy	595	AAAAACAATTAGAACCAACGTGCTACTAACCTTGAGGTTTGTCATCAACTGCTTATPAT	654
Db	4205	AAAAACTAAAAATTAACGAAAAAAATCTATATAAAAAAACATPATATTAATTAATTAAC	4146
Qy	655	CAAAATGGCAATTAATTATGTGATCTATA--CAATAAGCTATGTAATTAATACTAAA	712
Db	4145	TCATATATAAAAAAATATACAGATATATATCTATTTTAAAAATATAAAAATTAACGCAA	4086
Qy	713	CAGTAGATCCACTAAATGGGGAAAGCTTTAGATTCTAATGAGATTCTACAGTTATC	772
Db	4085	AAATTAACACATTAAACATATAAAAAACCTTTTATATATATATCTCCACCACTTAAAA	4026
Qy	773	GGAATATTATATATACGTTATCAACTATTAATGAACAAAGACTPATGCTGATGCATTAT	832
Db	4025	TTTTTAATAAAAATTAATATATCTCAACTTCAAACTTAAAAAT-ATATATTAATAACTT	3967
Qy	833	CTAATAGTTTATTAATAAAGTATCAAAATTAATGAACAA	874
Db	3966	CAAAAAATTTATACCAACATTAATTAATTAATTAATGA	3925

```

RESULT 14
US-10-311-455-1221/c
; Sequence 1221, Application US/10311455
; Publication No. US20030143606a1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Expression of Specific Genes
; TITLE OF INVENTION: Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1221
; LENGTH: 5163
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1221

```

[illegible]

Db	148	TATTATTTTATTTTATTTAGTCATCTTTAAGATATTAATATCTATATTTCTATGTAT	207
Db	4592	TCTCTCTCTATATATACATATATATTAANAATATTAATTAATATTTTATATTAATCTATA	4533
Qy	208	AAGAAAAGATCATCT---TAAGACTATTAAGTTGTTAGTGAACAATCTTTCTTAG	263
Db	4532	AATTCATATCATCTTAATTAATAAAATCTAACAAAAATTTTCATATAAAATCTAACAAATCTAA	4473
Qy	264	CATTGGGATTTCTACCTGTATGTCTATTAATAAAAGATGCAAAACCAATATATGGCCA	323
Db	4472	CMAATTAATTCATTCGCATTAATAAAATAAATTAACATTAATAAAATTAATTAATAAAATTTCTAATA	4413
Qy	324	AACCAATTTAGAGAGAGCGCAATGGAGTTAAACAGATCTATCATGATCTAAAGCATGAC	383
Db	4412	AAACTTAATTAATAAAATTAATAAAATTTACCTCAAAAATAATTAATAAAATCTTATATAAATATAC	4353
Qy	384	ATTAGCTTCACCTACAGAGCTATGCGAAGTTGAAGTGAAGTATTATCATCTGCTTATAGTGA	443
Db	4352	AATAATTAATAATATATATTAATTCACAAAAATTAATAATTAATAATCTTAATACTTAATAATTA	4293
Qy	444	AGCTGAAAACGTTAACATTAACCTTTATGCAATTAAGAACACTAATAAATNGCTAAANA-	502
Db	4292	ATTAATATCCAAATTAATAATGACAAACATATTAACAAACCTTTATATATTAATAAATAAACCATAAANA	4233
Qy	503	--CTAATTAGATCGAGCATCAACCAAGCTAATATACGATTAATAACGATTTGATTAATGA	560
Db	4232	ACAAAAAATAAACTTAATCTCTTTATATATATTAATAAATAATTAATCATTTCCATATANA	4173
Qy	561	ACACCCAAATTTAGTTGAAGCATACAAAGCACTTAATAACCACTTTAGAACAAACGTGTAC	620
Db	4172	AAAAATACCTTAATTAATTAACCTTA-ATACCATTAACAAATAATTAATCTAATAATTAATTA	4114
Qy	621	TAACTTGAAGTTGTGATCACTCACTGCTTATATCAAAATTCGCAATATTTAGTGATCT	680
Db	4113	AAATTTAATATTAATAAATTAATACTTCAAAAAAATAAATAAATAAACAATATTTTAACTT	4054
Qy	681	ATACATAAAGCTAGTATGTTATATACTAAACACATGATTCCTAATGGGGAACGCT	740
Db	4053	TTAAATTCAAAAAATAATTACTTTAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3994
Qy	741	TTTAGATT---CTAATGAGATTACTACAGTTAAATCGGAATATTAATATACGTTATCAAC	797
Db	3993	AATACCTTTTAAGCTATATCAAAATTTAAACCTCTTTATTAACAAAAACATAACATATTTAA	3934
Qy	798	TATTAATGAACAATAACATTAATGCTGATGATTTATCTAATAGTTTATTAATAAAGTAT	857
Db	3933	AAACCAACACACAAATTAATAAATTAATATATTAATTAATTAATTAATAAATAAATAAATAAT	3874
Qy	858	TCAAAAATATGAACAAGTTT	879
Db	3873	TCAAAAATATATTAATAAATTTCT	3852

```

RESULT 15
US-10-311-455-1797/C
; Sequence 1797, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424

```

```

; SEQ ID NO 1797
; LENGTH: 8771
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1797

```

```

Query Match          5.0%; Score 65.4; DB 12; Length 8771;
Best Local Similarity 48.2%; Pred. No.0.061;
Matches 284; Conservative 0; Mismatches 291; Indels 14; Gaps 3;

```

```

Qy 34 TTAACAAAAACACAAATCTTCTAACAAATCTTAATATAATAGCGCTTAATTAACTAA 93
Db 6791 TAAACAAAACTTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAAC 6732
Qy 94 AAAATTTAAAAAAGTGTCTTTCTTATCAACCAAAATCTCTAGTATTAACGCTATTYA 153
Db 6731 ACACAAAAATATATATATATATATATATATATATATATATATATATATATATAT 6672
Qy 154 TTTTATTTTATAGTCATCTTTTAAAGATATATATATATCT-----TATATCTATGA 205
Db 6671 TTCCAACTTAATACCAACCAACCAAAATATATATATATATATATATATATATATAT 6612
Qy 206 ATAGAAAAAGATCATCTTAAGACTATAGTTGTAGTACAAACATCTTTCTTAGCA 265
Db 6611 AAATTTTAAACATATTTTAAATAGCTAAATTAATTAATTAATTAATTAATTAATTA 6552
Qy 266 TTGGATTTCTAGCTGTATGTCTATTACTAAAAAAGATGCAACCCAAATATATGCCAA 325
Db 6551 CACTATATTTCTTTTATATATATATATATATATATATATATATATATATATATAT 6492
Qy 326 CCCAATTGAGCGCGCGGATGAGTGAATGATCTAATCAATGCTAAAGCATGACAT 385
Db 6491 TAAATTTACCAATATATATATATATATATATATATATATATATATATATATAT 6432
Qy 386 TA--GCTTCACTACAGACTATGCGCAAGATTGAAGCTAGTTATCATCTGCTTATAGTA 443
Db 6431 AAACCTTATATATATATATATATATATATATATATATATATATATATATATAT 6372
Qy 444 AGCTGAAAAGTTAACTTAATGCAACATTAGACACTTAAACATTAATATGCTAAAC 503
Db 6371 AAATATATTTATATATATATATATATATATATATATATATATATATATATAT 6316
Qy 504 TAAATTTAATCAGCATCAACCAAGCTAATACGATTAAGACTTTTGATTAATGACA 563
Db 6315 TAAACAAACAAACAGTACAAATATATATATATATATATATATATATATATAT 6256
Qy 564 CCCAAATTTAGTGAAGCATACAAAGACATAAAGACACTTAAAGCACTTAGACAA 612
Db 6255 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 6207

```

Search completed: August 25, 2003, 06:06:33  
Job time : 1251 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 03:07:11 ; Search time 3107 seconds  
(without alignments)  
10216.172 Million cell updates/sec

Title: US-09-901-572A-1  
Perfect score: 1306  
Sequence: 1 aaaaacatcagatgttcaat.....taaatcgtttatcaggtc 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: em\_estfun:\*  
15: em\_estcom:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_fun:\*  
21: em\_gss\_mam:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_pro:\*  
24: em\_gss\_rnd:\*  
25: em\_gss\_phg:\*  
26: em\_gss\_vrl:\*  
27: gb\_gss1:\*  
28: gb\_gss2:\*  
29: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.4	6.1	1101	29	CNS0039G
2	79.2	6.1	1200	13	AL063921 Drosophila
3	78	6.0	1124	13	BX437758 BX437758
4	77.2	5.9	1101	29	BX436282 BX436282
					AL064091 Drosophila

Result No.	Score	Query Match	Length	ID	Description
5	77.2	5.9	1200	13	BX437758
6	77.2	5.9	1201	13	BX446437
7	76.4	5.8	1201	9	AL565455
8	76.2	5.8	1101	29	CNS00EVL
9	75.8	5.8	1101	29	CNS000B8
10	75.6	5.8	1101	28	AQ946120
11	75.4	5.8	1225	9	CC238324
12	75	5.7	1201	9	AL536104
13	74.8	5.7	1201	29	CNS0167M
14	74.2	5.7	998	13	BX436885
15	74.2	5.7	1201	13	BX439779
16	73.4	5.6	1146	29	CNS021G2
17	73	5.6	1201	13	BX420717
18	71.8	5.5	1201	13	BX356896
19	71.2	5.5	960	13	BX346155
20	71	5.4	781	29	BX145762
21	71	5.4	1056	13	BX415058
22	71	5.4	1091	13	BX424950
23	70.6	5.4	1056	13	BX415058
24	69.4	5.3	1092	29	CNS020K7
25	68.8	5.3	1101	29	CNS00E07
26	68.8	5.3	1101	29	CNS016L1
27	68.6	5.3	1074	29	BZ686936
28	68.6	5.3	1201	13	BX433343
29	68.6	5.3	1201	13	BX361152
30	68.4	5.2	1187	9	AL514085
31	68.2	5.2	1098	13	BM803974
32	68.2	5.2	1200	13	BX414560
33	68	5.2	1178	29	BZ686021
34	67.8	5.2	759	29	CNS06QXV
35	67.8	5.2	1165	13	BX338369
36	67.8	5.2	1201	13	BX335216
37	67.8	5.2	1434	29	BZ576813
38	67.6	5.2	576	29	CNS035N7
39	67.6	5.2	1131	29	CNS034FO
40	67.4	5.2	994	13	BX414650
41	67.4	5.2	1101	29	CNS00EPO
42	67.4	5.2	1101	29	CNS016L1
43	67.2	5.1	1200	13	BX437739
44	66.8	5.1	839	13	BO151187
45	66.8	5.1	1101	29	CNS0021J

## ALIGNMENTS

RESULT 1  
CNS0039G 1101 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TERT end of BAC #  
DEFINITION BACROBK10 of RPL1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL063921 GI:4941778  
VERSION AL063921.1  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pretergata;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 1101)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT  
- Web : www.genoscope.cns.fr  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>  
The BDGP Drosophila melanogaster BAC library was prepared by Kazuoto Oseogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of



83 AAATTAACTAAATAATTAATAATAATGGTTTCTTATCAACCAAATTCCTAGTAATA 142

QY 83 AAATTACTTAAAAATTAAAAAATGGTTTTCCTATCAACCAAAATTCCTAGTAATTA 142

[illegible]

## COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazuhiro Oosawa and  
Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

## source

1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K08"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

BASE COUNT 395 a 120 c 103 g 334 t 149 others

## ORIGIN

Query Match 5.9%; Score 77.2; DB 29; Length 1101;  
Best Local Similarity 40.4%; Pred. No. 1.1;  
Matches 157; Conservative 47; Mismatches 185; Indels 0; Gaps 0;

QY 13 TTGTAATCTGATCTTTGCTTAAACCAAAATCTTACCAAAATCTTAAATA 72  
DB 664 TTATTTANWMAAATTAATAAAAAAAAAAAAAAAAAATTAATAAAATWTW 723  
QY 73 ATACCGCTTAATTAACCTAAATAAATGTTTCTTATCAACCAAAATCT 132  
DB 724 TAAATATATATATTAATAAATTAATAAATAAATAAATTAATTAATAAAT 783  
QY 133 CTAGTAATAAAGCTTATTTTATTTTATTTAGTATCTTTTAAAGATATAATATCT 192  
DB 784 TATTAATTAATATATATATATATATATTTTAAATTTTAAATTTAAATTTAAATTTN 843  
QY 133 TATATTTCTATGAATTAAGAAAGATCATCTTAAAGCATATGTTTGTAGTCAACA 252  
DB 844 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 903  
QY 253 TCCCTTTCTAGCATTTGGGATTTCTAGCTGATGCTATTAATAAATAAGATCAACCA 312  
DB 904 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 963  
QY 313 AATAATGGCAAAACCAATTTAGACAGCGCATGAGTAAACAGATCTAATCAATCT 372  
DB 964 AATATTTTAT 1023  
QY 373 AAAGCATGACATTTAGCTTCACTACAAGA 401  
DB 1024 AATAACMAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1052

## RESULT 5

## LOCUS

BX437758 1200 bp mRNA linear EST 15-MAY-2003

DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01

5-PRIME, mRNA sequence.

## ACCESSION

BX437758 BX437758

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1200)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Peng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com>, Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSOCAP008CA01QPL.

## FEATURES

## source

1. 1200  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOCAP008YB01"  
/clone\_lib="THYMUS"  
/clone\_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

BASE COUNT 515 a 30 c 71 g 310 t 274 others

## ORIGIN

Query Match 5.9%; Score 77.2; DB 13; Length 1200;  
Best Local Similarity 30.6%; Pred. No. 1.1;  
Matches 256; Conservative 156; Mismatches 418; Indels 6; Gaps 1;

QY 16 TTATCTGATATCTTTGCTTAAACCAAAATCTTCAACAAATCTTAATAATA 75  
DB 365 TTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 424  
QY 76 AGCGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 135  
DB 425 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 484  
QY 136 GTAATTAACGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 195  
DB 485 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 544  
QY 196 TATCTGATATTAAGAAAGATCATCTTAAAGCATATGTTTGTAGTCAACAATCC 255  
DB 545 TTTTAAATTAATTAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTTT 598  
QY 256 TTTCTAGCATTTGGGATTTTCTAGCTGATGCTATTAATAAATAAGATCAACCAAT 315  
DB 599 TAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 658  
QY 316 AATGGCAAAACCAATTTAGAAAGAGCGGAATGAGTAAACAGATCTAATCAATGCT 375  
DB 659 TAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 718  
QY 376 GCGATGACATTAAGTTCACTACAAGCATGCGCAAGTGAAGTAGTATTAATCACTGCT 435  
DB 719 AGRATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 778  
QY 436 TATAGTGAAGCTGAAGAGTAAACATTAACCTTAATCAATTAAGAACTAATAATG 495  
DB 779 TTTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 838  
QY 496 GCTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 555  
DB 839 AATAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 898  
QY 556 AATGAACCAACCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 615  
DB 899 AATAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 958  
QY 616 GCTACTAATCTTGAAGTTGTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 675

Db 221 AAAAAAAAAA WAAATTTTAAA 280

QY 327 CCAATTGAAACACGCGGATGAGTTTAAACAGATCTAATCAATCTAAAGCATGACATT 386

Db 281 AAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGMAAAACCMCCCAAAAAAAAAAC 340

QY 387 AGCTTCACTACAGACTATGCGCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTAAGC 446

Db 341 AACCAAAAAAAAAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAATWAAAAAAAAA 400

QY 447 TGAACAGTTTACATATACCTTAATGCACTTGAAACCACTAAATATGCTAAACCTTA 506

Db 401 AACCAATTT 460

QY 507 TTTAGAAATCAACCATCAACCAAGCTAAATACGATAAACGACTTTGATAATGAACACC 566

Db 461 AATTA 520

QY 567 AAATTGAGTTAGACATACAAAGCATTAACCACTTTAGACACAGCTGCTACTAACC 626

Db 521 AA CTTTTTCCCAAAWTTTTTTTWA 580

QY 627 TGAAGTTTGCATCACTGCTTAATTAATCAAAATGCGAAATTAATTAGATCTATCA 686

Db 581 WWWWTTTTTBAAAAADBRATTCYMAACCCCTCTWAAAAAAAAA CYYCAMEHYMAAAAA 640

QY 687 TAAAGCTAGTATGTTAAATTAACCTAAACACTGATGCCATAATGAGGGAACGCTTTTGA 746

Db 641 AAAAAAACCCCTTTTCTTTTAAATTAACCTTATGCGCHWMAACCTTTTBAACCCCTT 700

QY 747 TTTCAATGAGATTACTACAGTTAAATCGGAATATTAATTAATGCTTATCACTATTA 806

Db 701 TTTTCMCTCTCTTTTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 760

QY 807 ACAAAAGCTAATGCTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 866

Db 761 AAA 820

QY 867 TGAACAAAGTTTGTAGG 885

Db 821 AAAAAAAAAARADWTTTSTKG 839

RESULT 7

AL565455

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL565455 1201 bp mRNA linear EST 12-MAY-2003

AL565455 Homo sapiens PETA1 BRAIN Homo sapiens cDNA clone

CS0DF005Y018 3-PRIME, MRNA sequence.

AL565455

AL565455.2 GI:30549492

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

unpublished

On Feb 16, 2001 this sequence version replaced gi:12916448.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9232.f For more information about this cluster, see

http://www.genoscope.cns.fr/

cdg-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.





/mol\_type="genomic DNA"  
/strain="TREGU27/4 GUTat 10.1"  
/db\_xref="taxon:5691"  
/clone="Sheared DNA-46123"  
/clone\_1ib="Sheared DNA"  
/note="Vector: pUC18; Site 1: SmaI; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from a cloned population of  
Trypanosoma brucei (TREGU27/4 GUTat 10.1) was mechanically  
sheared to give a tight size distribution (approx 2 kb).  
The V + i method used for the library construction is  
described in detail in Smith, H.O. and Venter, J.C.  
(Making small insert libraries for whole genome shotgun  
sequencing projects. In Genome Sequencing: A Practical  
Approach, eds. M. Vaudin and B. Barrell, Oxford University  
Press, 1999)."

BASE COUNT 213 a 35 c 10 g 363 t  
ORIGIN

Query Match 5.8%; Score 75.6; DB 28; Length 641;  
Best Local Similarity 46.0%; Pred. No. 2.2;  
Matches 255; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

QY 750 TAATGAGATTACTACGTTAATCGGAATTTAATTAATGATCACTAATTAATGACA 809  
DB 624 TAGTAACTACTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 565  
QY 810 AAGAGCTAATGCGATGATTAATCTAATAGTTTAAATTAATTAATTAATTAATGA 869  
DB 564 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 505  
QY 870 ACAAGCTTTTGAAGGACTTTTCAAAAGCTTAATGTTCACTCAACTCAAGTTTGT 929  
DB 504 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 445  
QY 930 TCGTTTACGTCGATGTAACCCGCTCAATTAATTAATTAATTAATTAATTAATTA 989  
DB 444 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 385  
QY 990 TGGTGATGACCTTCAAGTAAGTAATCTTGCAACAGAAATGATGACAGATGTTCTTG 1049  
DB 384 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 325  
QY 1050 GATTTAATGTTAGCTGGAACAAACAGAGTACCAATTTAGTTTGAACATAGTGC 1109  
DB 324 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 265  
QY 1110 ATCAAGCTGTTATTTATTTCCCTAATAGTGTGTTAAAGCAGTGAATTAATAGT 1169  
DB 264 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 205  
QY 1170 TGGATTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1229  
DB 204 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 145  
QY 1230 TAGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1289  
DB 144 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85  
QY 1290 AATGCTTTTATCAG 1303  
DB 84 AAGAGCTTTTGAAG 71

RESULT 11  
CC238324 1225 bp DNA linear GSS 12-MAY-2003  
LOCUS CH261-19212.RM1.1 CH261 Gallus gallus genomic clone CH261-19212,  
DEFINITION genomic survey sequence.  
ACCESSION CC238324  
VERSION CC238324.1 GI:30564987  
KEYWORDS GSS.  
SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 1225)  
AUTHORS Kremitzki, C., Higgins, J., Wylie, K., Carter, J., McPherson, J.,  
Warren, W., Graves, T., Mardis, E. and Wilson, R.  
Gallus gallus BAC End Reads  
JOURNAL Unpublished  
COMMENT Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert length: 182000 Std Error: 0.00  
Seq primer: RM1 TAGCACTCACTAAGGAGA  
Class: BAC ends  
High quality sequence start: 44  
Location/Qualifiers  
1..1225

FEATURES  
source  
1..1225  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-19212"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/note="Vector: pTAKBAC2.1, Site 1: EcoRI; Site 2: EcoRI;  
CH261 Female Chicken library - for library and clone  
ordering information: http://www.chori.org/bacpac"

BASE COUNT 548 a 32 c 23 g 379 t 243 others  
ORIGIN

Query Match 5.8%; Score 75.4; DB 29; Length 1225;  
Best Local Similarity 38.8%; Pred. No. 1.8;  
Matches 330; Conservative 0; Mismatches 512; Indels 9; Gaps 2;

QY 4 AACATCAGATGTTAATCTGATATCTTGTCTTAAACAAACAAATCTTCAACAAAT 63  
DB 259 AATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 318  
QY 64 CCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 123  
DB 319 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 376  
QY 124 CAAATTTCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 183  
DB 377 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 436  
QY 184 AATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 243  
DB 437 AATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTA 496  
QY 244 GGTACAACTCTTCTTGAATTTGAGTTCTGATGATGATGATGATGATGATGATGAT 303  
DB 497 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 556  
QY 304 GGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 363  
DB 557 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 616  
QY 364 ATCAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 423  
DB 617 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 676  
QY 424 TTAATCATCTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 483  
DB 677 TTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 736  
QY 484 CAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 543  
DB 737 NATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 796



Db		1184 ATATATAAAYACGTAAMWTTAAATATATAAAATHTWTAAAMAAMTWTAATATAATACAAMWWMTA	1125
Oy		109 GTTTTTCTTNTCAACCAAANTCCTCGTAGTAATAAGCCATTATTTATTTTATTTTGTC	168
Db		1124 TWHHTATTTTTTTTTTTTTTWMATAAATATMWTMTWTTTTTYAAAAAMAATMWTTTTTAAMA	1065E
Oy		169 ATCTTTTAAGATATAAATATATCTTAATATTTCTATGTAGATAGAAGAAGATCATCTTAAG	228
Db		1064 WTATWTMATTTAAAANAAMATWMTTTTTTATWTATSMATTWMAAANMAAATAATMATAAW	1005F
Oy		229 ACTATAGTTGGTTAGTGACAAACATCCTTCTTGACATGGGATTTCTAGCTGATGCT	288
Db		1004 WATATTTWMTATWTTWTTTTTAAATAAATAWTTWTTMMWTTATTTTWCWTATATATATTW	945
Oy		289 ATTACTAAAANAAGATGCAACCACCAATTAATGGCCAACCACAATTAGAAGCGCGGATG	348
Db		944 TWMTWTMACCTCAATWAATTTMMWMAAATAWTTWTAABAAWTTATTAATATAAAMATTAAWTT	885
Oy		349 GAGTTAACAGATCTAATCAATGCCTAAGCGATGACATTAGCTCACCTACAGACATATGCC	408
Db		884 AAATTTAAAANAATPAATTTTATWBAABAATRWMAAATPTWAAAAMWTTTWTATWATAATMA	825
Oy		409 AAGATTGAAGCTAGTTATCATCTGCTTTATATGTGAAGTGAAACAGTTAACATPACCTT	468
Db		824 AAAAMTTTMAWATTTWATTAATTAATWMAAANAAMWMAAATAWATWYATTAATATAA--A	768
Oy		469 AATGCACATTTAGAACCAATAAAATGGCTAAATACCTAATTAGATGCACCATCAACCA	528
Db		767 AATTTTAAAATATATWAAAMWTTATATATWMAAAMAAATTTATATATTTATTTWATATAA	708
Oy		529 GCTAATACGATTAACCAACGACTTTTGATATATGACACCAAAATTTAGTTGAA	579
Db		707 ATAAMAAAATAATATAAAMATTAAMAAATTTTATATATAHATTATATAAMWHA	657
RESULT_13	CNS0167M/c		
LOCUS			
DEFINITION	CNS0167M	1201 bp DNA linear	GSS 26-JUL-1999
ACCESSION	BACN15M24	Drosophila melanogaster genome survey sequence T7 end of BAC	
VERSION	ALI06396	BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
KEYWORDS	ALI06396.1 GI:5621701	GSS.	
SOURCE		Drosophila melanogaster (fruit fly)	
ORGANISM		Drosophila melanogaster	
REFERENCE		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS		1 (bases 1 to 1201)	
TITLE		Genoscope.	
JOURNAL		Direct Submission	
COMMENT		Submitted (22-JUL-1999) Genoscope - Centre National de Sequencage BP 191 J1006 EVRY cedex - FRANCE (E-mail : seqet@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (BDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Bilaud at CPHF (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.	
FEATURES	source	location/Qualifiers	
		1..1201	
		/organism="Drosophila melanogaster"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:7227"	
		/clone="BACN15M24"	
		/clone_lib="DrosBAC"	
		/plasmid="pBelobAC11"	
		/note="end = T7"	

[illegible]

```

source
1. .998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP004YC09"
/close_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT      593 a      98 c      56 g      157 t      94 others

ORIGIN

Query Match      5.7%; Score 74.2; DB 13; Length 998;
Best Local Similarity 42.9%; Pred. No. 2.7;
Matches 249; Conservative 37; Mismatches 290; Indels 4; Gaps 2;

QY      34      TTAACCAACCAAAATCTTCTAACAACAAATCTTAATAATTAAGCCGTTAATTA 93
DB      384      TAAAAAATTAATAATTAATAATTAATAATAATAATAATAATAATAATAATAATA 443

QY      94      AAAATTAATAAATGGTTTCTTATCAACCAAAATCTAGATATTAACGCTTAATTA 153
DB      444      AAAATTAATAAATCGTTTCTTATCAACCAAAATCTAGATATTAACGCTTAATTA 503

QY      154      TTTTATTTTATGATCTTTTATAGATATTAATAATATCTTAATATTTCTATGATA 213
DB      504      WAAAAATTAATTTTATTAATAAAMWMAAAATATTAATAAATAATTAATAAATA 563

QY      214      AGAATCATCTTAAGAATTAATTAAGTTGTGTGATGACATCCCTTTAGCATGGGAT 273
DB      564      AAAAAAATAAATAAATAAATAATTTATCAAAAAAATAAATAAATAAATAATTA 623

QY      274      TTAGCTGTATGTCTATTTACTTAATAAAGATGCAACCAATAATTAAGCCAAATTA 333
DB      624      AAAATTAATAAATCTTCTTAAATAAAGATGCAACCAATAATTAAGCCAAATTA 680

QY      334      GAAAGCAGCGGAATGAGATTACAGATCTAATCAATGCTAAAGGATGACATTAGCT 393
DB      681      AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 740

QY      394      CTCAAGACTATGCCAAGATTGAAGTAGTTTATCATCTGCTTATAGGAAGTGAACA 453
DB      741      AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 800

QY      454      GTTA-ACATAATACCTTAATGCAACATTAGAAACAATAAATGCGCTAAATAATTGA 512
DB      801      TATCAAAATMTATTTNAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 860

QY      513      ATGAGCATCAACCAAGCTAATAGCATTAACGACTTTGATATGAAACCCCAATTT 572
DB      861      AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 920

QY      573      AGTGAAGCATACAAGCACTAATAAACCACTTTAGAACAA 612
DB      921      CTTTCWATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960

RESULT 15
LOCUS      BX439779      1201 bp      mRNA      linear      EST 15-MAY-2003
DEFINITION BX439779. Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE014YF00.
ACCESSION  BX439779
VERSION     BX439779.1 GI:30771778
KEYWORDS    3-PRIME, mRNA sequence.
SOURCE      EST.
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1201)

```



THIS PAGE BLANK (USP10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using BW model

Run on: August 14, 2003, 10:11:57 ; Search time 25 Seconds  
(without alignments)  
857.767 Million cell updates/sec

Title: US-09-901-572a-3

Perfect score: 2324  
Sequence: 1 MHYFRNCIFPLVILYGTN.....SSNENNADKIPGYRRPGTFL 456

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	314	13.5	865	1	VGLB_HSVMD
2	151	6.5	2660	1	YEBJ_ECO57
3	150.5	6.5	682	1	NISP_LACIA
4	149.5	6.4	857	1	ARS6_CANAL
5	147.5	6.3	1085	1	CUT7_SCHRO
6	142	6.1	719	1	VM41_YEAST
7	141	6.1	1314	1	SW11_YEAST
8	140	6.0	3178	1	Y589_CAEEL
9	138	5.9	1260	1	AL51_CANAL
10	137.5	5.9	1018	1	FNBA_STANU
11	137	5.9	1010	1	YK11_CAEEL
12	136	5.9	995	1	Y109_YEAST
13	134.5	5.8	704	1	MSN2_YEAST
14	134	5.8	444	1	PS11_YEAST
15	134	5.8	1140	1	YV96_YEAST
16	134	5.8	1271	1	Y338_MYCE
17	133.5	5.7	956	1	YEP3_YEAST
18	133.5	5.7	1161	1	DAN4_YEAST
19	133	5.7	1381	1	YBE7_YEAST
20	132.5	5.7	4164	1	BAG_STRAG
21	130	5.6	1251	1	YBVO_YEAST
22	129.5	5.6	903	1	VGLB_HSVIF
23	129.5	5.6	904	1	VGLB_HSV11
24	129.5	5.6	904	1	VGLB_HSV1P
25	129.5	5.6	1117	1	YV96_YEAST
26	129.5	5.6	1159	1	N124_SCHPO
27	129.5	5.6	1251	1	P2P_LACIC
28	129	5.6	1251	1	RBP2_PLAVB
29	128.5	5.5	750	1	YK57_YEAST
30	128.5	5.5	2358	1	YEBJ_ECO11
31	128	5.5	719	1	PBP4_STR6
32	128	5.5	837	1	R0D1_YEAST
33	128	5.5	1433	1	CAT8_YEAST

34	128	5.5	1902	1	P3P_LACIC	P15222 lactococcus
35	127.5	5.5	544	1	FLGR_BUCAP	Q6K90 buchnera ap
36	127.5	5.5	800	1	INLA_LISMO	P5146 listeria mo
37	127.5	5.5	1858	1	P3K2_DICDI	P54674 dictyostell
38	127.5	5.5	3712	1	LMA_DROME	000174 dipterosphila
39	127	5.5	537	1	ARP_PLAPA	P04931 plasmodium
40	127	5.5	866	1	MYSP_SCHJA	Q05870 schistosoma
41	127	5.5	979	1	P115_MYCHR	P11508 mycoplasma
42	127	5.5	1142	1	GIN4_YEAST	Q12263 sacccharomyc
43	126.5	5.4	1672	1	PMPB_CHLMU	09P1Y2 chlamydia m
44	126.5	5.4	2334	1	WAPA_BACSU	007833 bacillus su
45	126	5.4	719	1	PBP4_STRPN	Q04707 streptococc

## ALIGNMENTS

RESULT 1	ID	VGLB_HSVMD	STANDARD;	PRT;	865 AA.
AC	P18538				
DT	01-NOV-1990	(Rel. 16, Created)			
DT	01-NOV-1990	(Rel. 16, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.				
GN	GB.				
OS	Marek's disease herpesvirus (strain RB-1B) (MDHV).				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphan herpesvirinae; Marek's disease-like viruses.				
OX	NCBI_TaxID=33707;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEBLINE-89293086; PubMed=2544666;				
RA	Ross L.J.N., Sanderson M., Scott S.D., Blinn M.M., Doel T., Milne B.;				
RT	"Nucleotide sequence and characterization of the Marek's disease				
RL	virus homologue of glycoprotein B of herpes simplex virus.";				
CC	J. Gen. Virol. 70:1789-1804(1989).				
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.ebi.ac.uk/announcements">http://www.ebi.ac.uk/announcements/</a>				
CC	or send an email to <a href="mailto:license@ebi.ac.uk">license@ebi.ac.uk</a> ).				
CC	-----				
DR	EMBL; D13713; BAA02866.1; -				
DR	InterPro; IPR000234; Glycoprot B.				
DR	Pfam; PF00606; Glycoprotein B; 1.				
DR	ProDom; PD000693; Glycoprot B; 1.				
KW	Signal; Glycoprotein; Transmembrane				
FT	SIGNAL	1	21		POTENTIAL.
FT	CHAIN	22	865		GLYCOPROTEIN B.
FT	DOMAIN	22	682		EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	683	700		POTENTIAL.
FT	TRANSMEM	709	729		POTENTIAL.
FT	DOMAIN	732	752		POTENTIAL.
FT	CARBOHYD	753	865		CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	184	184		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	332	332		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	364	364		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	406	406		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	425	425		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	631	631		N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	865 AA;	98091 MW;	B30E93C1AC65C6C3	CRC64;

Query Match 13.5%; Score 314; DB 1; Length 865;

Best Local Similarity 100.0%; Pred. No. 7.2e-10; Mismatches 0; Gaps 0;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MHYFRNCIFPLVILYGTNSSPTQVNTSREVSSVQASBEESTFLCPVPVSTVIRL 60

DB 1 MHPFRNCFPLVILVGTGSSPSTQNTVTSREVSVOLSEESTFYLCPPVGVSTVRL 60  
QY 61 E 61  
DB 61 E 61

RESULT 2  
ID YEED ECO57 STANDARD; PRT; 2660 AA.  
AC Q8X8V7; Q8X2B9; Q8X2C0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein yeed.  
GN Z3135 OR ECS2775/ECS2776.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
CX NCBI\_Taxid=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postel G., Hackett J., Link S., Boutin A., Shao Y., Miller D.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"  
RL Nature 409:529-533 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takai H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22 (2001).  
CC -1- SIMILARITY: Contains 16 Big-1 domains.  
CC -1- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.  
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
CC frameshift in position 1315.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DB EMBL, AB005423; AAG57041.1; -  
DB EMBL, AP002559; BAB36198.1; ALT\_FRAME.  
DB EMBL, AP002559; BAB36199.1; ALT\_FRAME.  
DR InterPro; IPR003344; Big-1.  
DR InterPro; IPR003345; Intimin.  
DR InterPro; IPR006601; PKD domain.  
DR Pfam; PF02369; Big-1; 16.  
DR PRINTS; PRO1369; INTIMIN.  
DR SMART; SMO0634; PKD\_1; 16.  
DR SMART; SMO0089; BID; 8.  
DB Hypothetical protein; Repeat; Complete proteome.  
FT DOMAIN 738 834 BIG-1 1.  
FT 840 929 BIG-1 2.  
FT DOMAIN 931 1033 BIG-1 3.  
FT DOMAIN 1042 1132 BIG-1 4.  
FT DOMAIN 1134 1236 BIG-1 5.  
FT DOMAIN 1245 1335 BIG-1 6.

FT DOMAIN 1337 1439 BIG-1 7.  
FT 1448 1539 BIG-1 8.  
FT DOMAIN 1548 1632 BIG-1 9.  
FT DOMAIN 1653 1750 BIG-1 10.  
FT DOMAIN 1751 1855 BIG-1 11.  
FT DOMAIN 1856 1957 BIG-1 12.  
FT DOMAIN 1963 2056 BIG-1 13.  
FT DOMAIN 2065 2156 BIG-1 14.  
FT DOMAIN 2157 2252 BIG-1 15.  
FT DOMAIN 2254 2355 BIG-1 16.  
SQ SEQUENCE 2660 AA; 280062 MW; 01EB92A08F5C09D2 CRC64;  
Query Match 6.5%; Score 151; DB 1; Length 2660;  
Best Local Similarity 23.0%; Pred. No. 1.8;  
Matches 108; Conservative 70; Mismatches 188; Indels 104; Gaps 20;  
QY 12 LIVILVGTGSSPSTQNTVTSREVS---SVQLS---EESTFYLCPPVGVSTVIRLEFGC 64.  
DB 1131 IVNIAPDASNAQVTLNIPAOQVVTNNSDSVOLFTATVDPNNH---PVAGITV----- 1179  
QY 65 MSTIKDANPNNGQTOLEAARMELTDLINKAMTLASLDQYAKIEASLSAYSAAETVNN 124  
DB 1180 -----NFTMPQVVAANFTLEN--NGIATIQANGEAHVTLKGGKAGTHVTATLGN 1227  
QY 125 NLNATLEQLMAKTNTLESAINQANT-----DKTT-----FDNEHPVLVAYK 166  
DB 1228 N-NASDAQPTFPAVDKDSAVVLOTSAEIIIGNVDITTLATVYKDFDN-----AVK 1279  
QY 167 ALKTTLEQRAITLGLSTAVYNOIRNLVDLYNKASSLIRKTLDPNLNGTILL--DSNEIT 224  
DB 1280 DLQVTF---STN---PADTQLSGKSNWTD-----SGVAETFK---GTVLGVHTAEAT 1324  
QY 225 TVNRNINNTLSTNEOKTNADALNSFIKKVIONNEGSPFGTFNNAVQSPNSFVAFSA 284  
DB 1325 LPNNNDNTKIVNIAPDASNAQVTLNIPAOQVVTNNSDSVOLATVYD--PSNHPVAGITV 1382  
QY 285 DVEFVNYKYARRFTWNGDEPSSRIANTNSITDVSMTYISLAGNTKYQSPSNYPSTGY 344  
DB 1383 NFTMPQVVAANFTL---ENNGLAITQANGEAHVTLGKKAQHT--VTATLSNNTSDSQ 1437  
QY 345 LYFPIKLVKADANNVGLQYKLNNGVQVYEPAT-----STSAMN 384  
DB 1438 ---PVEFVADKTSALVVLQISKEITNGVDSALTATVYDQEPNEVNNLPVTFSTASG 1494  
QY 385 TTANPTPA-VDEIVAKIVISGLRFGQNTLESPYTERGMNKYAPMIGN 433  
DB 1495 LTLPGESNTGESGIQAOTLAGVAFGEQIVTASLANNGASDNKTVHETGD 1544

RESULT 3  
ID NISP\_LACLA STANDARD; PRT; 682 AA.  
AC Q007596;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Nisin leader peptide processing serine protease nisp precursor  
DE (EC 3.4.21.-).  
GN NISP.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
CX NCBI\_Taxid=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NIZO RS;  
RX MEDLINE=93239683; PubMed=8478324;  
RA van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,  
RA Kuipers O.P., de Vos W.M.;  
RT "Characterization of the Lactococcus lactis nisin A operon genes  
RT nisp, encoding a subtilisin-like serine protease involved in  
RT precursor processing, and nispR, encoding a regulatory protein  
RT involved in nisin biosynthesis.";  
RT J. Bacteriol. 175:2578-2588 (1993).

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6F3;
RA MEDLINE=94213458; PubMed=8161176;
RA Engelke G., Gutowski-Eckel Z., Kieseau P., Siegers K.,
RA Hammeleimann M., Entian K.-D.;
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
RT 6F3.";
RL Appl. Environ. Microbiol. 60:814-825(1994).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=95357326; PubMed=7630881;
RX Stizen R.J., Rolfe H.S., Kuipers O.P., de Vos W.M.;
RT "Homology modelling of the lactococcus lactis leader peptidase Nisp
RT and its interaction with the precursor of the lantibiotic nisin.";
RL Protein Eng. 8:117-125(1995).
CC -1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC -1- PATHWAY: Nisin biosynthesis; last step.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11061; AAA5200.1; -
DR EMBL; X76884; CAA54210.1; -
DR PIR; S44131; S44131.
DR HSSP; P29600; 1GCI.
DR MEROPS; S08.059; -.
DR InterPro; IPR006192; LPTXG.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; FALSE NEG.
DR PROSITE; PS50847; GRAM POS ANCHORING; FALSE NEG.
KW Hydrolyase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW Signal.
FT SIGNAL. 1 22 POTENTIAL.
FT PROPEP 23 195 POTENTIAL.
FT CHAIN 196 655 NISIN LEADER PEPTIDE PROCESSING SERINE
FT PROPEP 196 655 PROTEASE NISP.
FT PROPEP 196 655 REMOVED BY SORTASE (BY SIMILARITY).
FT ACT_SITE 259 682 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 306 306 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 512 512 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 652 656 LPTXG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 655 655 AMIDE-LINKED TO CELL WALL (POTENTIAL).
FT CONFLICT 500 500 A -> T (IN REF. 2).
SQ SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Query Match 6.5%; Score 150.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.34;
Matches 75; Conservative 60; Mismatches 131; Indels 91; Gaps 13;

```

```

QY 299 -----WNGDEPESRLANTNSITDVSWIYSLAGTNTKYPQSFSENGPS 341
DB 194 LRQPLKQKVEAQPILNLSNSEKASVYTNSHDFWDQW-----DMKY---VTNNGES 243
QY 342 TGVLYPPKLVKADANNVGLQYLANGNVQVAFSTSN-----NT 385
DB 244 VA-LYOPSKKI-----SVGI---IDSGIMEHPDLNSLGNVPEKNLVPKCGFNDPEPDE 293
QY 386 TANPTPAVDEIKAKIVLSGLRFGQNTIELSVPEGGMNKVAMIG-NITLSSNEN 441
DB 294 TGNPSDIVD-----KMGHGTVEAQQITANGNIIIGAVGITVNIYRVEGEN 338

RESULT 4
AR56 CANAL
ID AR56 CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ARG5,6 protein, mitochondrial precursor [contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5,6
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gali C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -1- PATHWAY: Arginine biosynthesis; second step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98880; CAA67383.1; -
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; ACGLUKinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinaase; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdehyde_dh; 1.
DR Pfam; PF02774; Semialdehyde_dhc; 1.
DR Pfam; PF001765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARGK; 1.
KW Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;

```

RX MEDLINE:184846401; PubMed:118593360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald J., McLean J.,  
 RA Mooney P., Mounle S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson S., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voicakeet G., Aert R., Robben J., Gymnopoulos B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,  
 RA Sipakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RL "The genome sequence of *Schistosoma* *mansoni*.";   
 CC Nature 415:871-880(2002).  
 CC -I- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM  
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES,  
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO  
 CC FORM A SHORT SPINDLE THAT ELONGATES TOWARD THE NUCLEUS AT  
 CC METAPHASE. MUTATIONS AT CUTY BLOCK SPINDLE FORMATION.  
 CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIOC  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X57513; CAA40738.1; -;  
 DR EMBL; Z70691; CAA94636.1; -;  
 DR PIR; T38378; T38378.  
 DR HSSP; P17119; 3XAR.  
 DR GeneDB Sprobe: SPAC25G10.07c; -;  
 DR InterPro: IPR0010752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin.1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.  
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN; 1.  
 KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;  
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.  
 FT DOMAIN 70 435  
 FT KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 436 604  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 715 740  
 FT COILED COIL (POTENTIAL).  
 FT NP BIND 159 166  
 FT COILED COIL (POTENTIAL).  
 FT REPEAT 987 998  
 FT ATP (BY SIMILARITY).  
 FT MOD RES 1011 1011  
 FT PHOSPHORYLATION (BY CDC2) (BY  
 FT SIMILARITY).  
 FT SASNPKRRRPPITDGYGDRSDTNSPT ->  
 FT LRAIINDVSLILTL (IN REF. 1).  
 FT  
 SQ SEQUENCE 1085 AA; 12213 MM; 5669277875559D58 CRC64;  
 Query Match 6.3%; Score 147.5; DB 1; Length 1085;  
 Best Local Similarity 20.0%; Pired. No. 0.88;  
 Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;



```

QY 17 YGINSSESTQNT-----SREVVSV-----QLSEESTFYLCPPVGSYVIR 59
Db 554 YETNEAKITTVATLDSQYRESKEYIASLYEKLDRTERRNNKNNNNFM-----N 602
QY 60 LEFGCMSTTK---KDNPNNGQ--TQLEAAMELTDLINAKA-MTLASL---QDYAKIE 109
Db 603 LKFNILMTLRSPHSGFTDETGYFTLLDNFNASMBELINTSNOLLISMTKITEHPOLD 662
QY 110 ASLSAYSEARTVNNNNATLEOLKMAKTLESAINOANTKTFTDNEHPN-----LYEA 164
Db 663 EALDSASASSCAVPNSLDLYSELKDSKSLDLALHSLQDISMSQGLGNGISSELTIEL 722
QY 165 YKALKTT---LEQATNLEGLSSPAYNOIRNNLVLYNKASLLITKTLDPUNGTLDSN 221
Db 723 QDKMKEYSLVQELVRLSYNLQHTHEESQKELMGVRDIDALVYCTTSLDADIDILSD 782
QY 222 EIT-----TVNRNINNTLSTNEQ-KTNADL-----SNFPIK--VI 256
Db 783 YISQKSKFESKQODLIANIKRIVSNFLQEQNESLYTQADILHSLNDTNSNIRKAEIM 842
QY 257 QNNBOSFVGTFTNANVPQSNYSFVAFSADYTPVNYKARFTVWG---DEPSRILANT 312
Db 843 NNRESEFL---RNA-----ASQAEIVGANKERIQKTVENGSQGLDSKSAIHSNS 889
QY 313 NSITDVSWIYSLA-----GTNTKYQ-----FSFSNYGPGSTGYLYPPYKLVKA 354
Db 890 RSMVD---HCLALAESQKQGVNLEVOITDLRLQKVKHSESDNTKEKHQOL--LDLLES 942
QY 355 ADANNVGL--QYKLANGVQVE---FATISANNTT-----ANFTPAVDELKV 398
Db 943 LVGNNDNIDSIKTPHTELQKITDHVLKGTISLANHTMELLGLGESLCNLETTEDTSL 1002
QY 399 AKIVLSGLRFGQNTI-----ELSVPTGEGNNKAVAMIGNIYLSNENNADKIPG 448
Db 1003 VKLETTGTPTSKRELPAFTPSWTRDSSLIKETTNLMDSDKFAVETIYSSNQNEPDV-- 1060
QY 449 YRRP 452
Db 1061 YDKP 1064

RESULT 6
Y41_YEAST STANDARD; PRT; 719 AA.
ID Y41_YEAST STANDARD; PRT; 719 AA.
AC 003213;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 79.4 kDa protein in Ald2-DDR48 intergenic region.
GN YMR172W OR YMR010.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Desman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsels K., Iye G., Moulé S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XII."
RT Nature 387:90-93 (1997).
RL -1 SIMILARITY: LOW, TO YEAST MSN1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/249808; CAAB9905.1; -
DR PIR; S55119; S55119.
DR TRASNPRAC; TQ4601; -.
DR SGD; S0004783; HOTI.
DR GO; GO:0000228; C:nuclear chromosome, IDA.
DR GO; GO:0006972; P:hyperosmotic response, IGI.
DR Hypothetical protein.
KW SEQUENCE 719 AA; 79415 MW; 4652BE93743D5A54 CRC64;

Query Match 6.18; Score 142; DB 1; Length 719;
Best Local Similarity 21.28; Pred. No. 1;
Matches 100; Conservative 69; Mismatches 175; Indels 128; Gaps 20;

QY 19 TNSPSTQNTSREVVSVVOLSEESTFYLCPPVGSYVIRLERGCMSTTKDNPNNGQ 78
Db 150 TNSPS--NEISTQQLKIFQMDMS-----ARMTMESSEFNKSNKTAEQNTMV 197
QY 79 TQLEAAMELTDLINAKAMTLA-----SLQDYAKIE-----ASLSAYSEARTVNN 124
Db 198 LNLKQDYKVMKNTLILKLVADPSARPSTNNQNKALIELNSISAVSAIYLOKQNNQ 257
QY 125 NLAATLEQLKMAKTLESAINQ-----ANTDKTTFDNEHPNLEAVKALKTILEQR 175
Db 258 SGRQHTADLCTGDSNTHSGINQHRRTTGITIDVNTTAQLNNQFSN-----ALNTILPDQ 311
QY 176 ATNLEGLSSPAYNOIRNNLVLYNKASLLITKTLDPUNGTLDSNITTVNRNINNTL- 234
Db 312 QHN-----RNVSSQNTNQ--SLPNRLQGVIN-----TANQOQSOVL 348
QY 235 -STINQKTNADALS--NSFIKKYIQNNEQSFV---GFTTNANVPQSNYSFVA---FSA 284
Db 349 HNTTTHQVNRSPISFNASTDQKPKLNPNGIKRRRNTQSNNASTINDASAAPQKISA 408
QY 285 DVTTPVNYKARFTVWNGDEPS--SRILANTNSITDVSWIYSLAGTNTKYQFSFSGYPPST 342
Db 409 -LSPLTNSHNSITSMYNTNSSIHSGVTSASNSFFHLN-----SLNPFSTTT 453
QY 343 GYLFPYKLVKADANNVGLQYKLANGVQVEPATISANNTTANFTPAVDELKVAIV 402
Db 454 ALSPLSLALDNASPPQNVIPPIINTQQLPSQILINDSTTS----- 498
QY 403 LSGLRFGQNTIELSVPTGEGNNKAVAMIGNIYLSNENNADKIPGRPRT 454
Db 499 -----EL-LPSGKSGVNT-----NIV---NRNPSLTLPSPKXMT 529

RESULT 7
SW1_YEAST STANDARD; PRT; 1314 AA.
ID SW1_YEAST STANDARD; PRT; 1314 AA.
AC P09547;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription regulatory protein SW1 (SWI/SNF complex component SW11)
DE (transcription regulatory protein ADR6) (regulatory protein GAN3).
GN ADR6 OR SW11 OR GAN3 OR YF0106W OR LPA1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=9169875;
RA O'Hara P.J., Horowitz H., Eichinger H., Young E.T.;
RA "The yeast ADR6 gene encodes homopolymeric amino acid sequences and a
RT potential metal-binding domain."
RL Nucleic Acids Res. 16:10153-10170 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=9169875;

```

Db	112	FTSPSTAVVNNNAAPANTVGKAKANFIQONSPQFNSPYDNNSSN	-----	TNLSLSFQAI-L	165
Qy	190	IRNNVLVPLYNKASGLTIKTLDPINGCGLLDBSNELTTYNRNINNTLSTINECKTADALSN	249		
Db	166	AKNSIIID-SSNLPLOAQOQLYGG-----NNNNNSTGIANDVITPHPTTN	209		
Qy	250	SFIKVIQNNNEQSFVGFFTNANVOP-SNYSFVAFSADVTPVNRYKYARBTWNCGDEPSSRI	308		
Db	210	--VQSIQSSSSSTP--TNSNSTPNNNQOLFPNNASN-----NGMLTSNQL	254		
Qy	309	LANTNSITDVSWIYSLAGTNTKTYPFSNSNGPSPGTYLPFKLYKADANVGLQYKL--	366		
Db	255	ISN-----YAASNSMDRSSASANEFPVNT-----SDNNNSNNNNHNRN	292		
Qy	367	-----NNGVQOQVEFATSTSANNTANPPPADE	395		
Db	293	NSNNKTSNNNNVTAVPAATPAINTNNSISNNTVASE	328		
RESULT 8					
YS89	CABEL	STANDARD;	PRT;	3178	AA.
ID	YS89	CABEL	009624;	009625;	Q669D4;
AC	01-NOV-1995	(Rel. 32, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DN	Hypothetical protein ZK945.9	in chromosome II.			
GN	ZK945.9/ZK945.10.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;				
OC	Rhabditidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Briscot N2;				
RA	Wilkinson-Sproat J.;				
RL	Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	REVISED.				
RA	Durbin R.;				
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- SIMILARITY: Contains 1 GPS domain.				
CC	-1- SIMILARITY: Contains 1 PLAT domain.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration -				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL; Z48544; CAB70192.1; -				
DR	EMBL; Z48582; CAB70192.1; JOINED.				
DR	EMBL; Z48582; CAB70201.1; -				
DR	EMBL; Z48544; CAB70201.1; JOINED.				
DR	WormReg; ZK945.9; CE25697.				
DR	InterPro; IPR002111; Cat_channel_TyPL.				
DR	InterPro; IPR005821; Ion_trans.				
DR	InterPro; IPR001024; Lipoxxygenase_LH2.				
DR	InterPro; IPR003915; PKD 2.				
DR	InterPro; IPR00203; PKD_cys_rich.				
DR	Pfam; PR01825; GPS; 1.				
DR	Pfam; PF00520; Ion_trans; 1.				
DR	Pfam; PF01477; PLAT; 1.				
DR	PRINTS; PR01433; POLYCYSTIN2.				
DR	SMART; SM00303; GPS; 1.				
DR	SMART; SM00308; LH2; 1.				
DR	PROSITE; PSS0095; PLAT; 1.				
DR	Hypothetical protein; Transmembrane.				
FT	TRANSMEM	13	30	POTENTIAL.	
FT	TRANSMEM	51	73	POTENTIAL.	

```

MoJ. Microbiol. 15:39-54(1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -1- PM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC -1- SIMILARITY: TO YEAST SAG1.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.emb-bib.ch/announce/
CC or send an email to license@isb-bib.ch).
-----
DR EMBL; L25902; AAC1649.2; -;
KW Cell adhesion; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 1260 AGGLUTININ-LIKE PROTEIN 1.
FT DOMAIN 433 792 10 X 36 AA TANDEM REPEATS.
FT REPEAT 433 468 1-1.
FT REPEAT 469 504 1-2.
FT REPEAT 505 540 1-3.
FT REPEAT 541 576 1-4.
FT REPEAT 577 612 1-5.
FT REPEAT 613 648 1-6.
FT REPEAT 649 684 1-7.
FT REPEAT 685 720 1-8.
FT REPEAT 721 756 1-9.
FT REPEAT 757 792 1-10.
FT DOMAIN 983 1152 2 X 26 AA APPROXIMATE REPEATS.
FT REPEAT 983 1043 2-1.
FT REPEAT 1092 1152 2-2.
FT DOMAIN 399 404 POLY-THR.
FT DOMAIN 408 418 POLY-THR.
FT DOMAIN 450 455 POLY-THR.
FT DOMAIN 486 491 POLY-THR.
FT DOMAIN 522 527 POLY-THR.
FT DOMAIN 558 563 POLY-THR.
FT DOMAIN 594 599 POLY-THR.
FT DOMAIN 630 635 POLY-THR.
FT DOMAIN 666 671 POLY-THR.
FT DOMAIN 702 707 POLY-THR.
FT DOMAIN 738 743 POLY-THR.
FT DOMAIN 774 779 POLY-THR.
FT DOMAIN 874 877 POLY-SER.
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 687 687 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 723 723 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 886 886 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 918 918 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;
-----
Query Match 5.9%; Score 138; DB 1; Length 1260;
Best Local Similarity 21.2%; Pred. No. 3.5;
Matches 103; Conservative 75; Mismatches 189; Indels 118; Gaps 21;
OY GNNS-----SPSTQWTSREVVSSVQLSEESTFYLCPPVGSTVIRLEFGCMSITKKDAN 73
Db 748 GTDIVIIRPPNPVITTEYWSQ---SFATTIVTAPPGGTIVIIIE--SMSSKISTKS 802
OY 74 PN-----NGQTQ---LEARMETLDLINAAMKAMTLAGLQYAKIEASLSS 114
Db 803 SNDITSIIPFSRPHVYNSTSDLPSTRESSMNTPTSISSDKMLSS-----TT 851
OY 115 AYSSEAIVNNNNLNTLEQLQMAKTNLESAINQANTDKTTDNEHPNIVEAYKALKTTLEQ 174
Db 852 LVYESETTTESICSDGKEC-----SRLLSSSGIIVTNPDSNESSIYVTSVTVPASTMSD 903

```

QY 175 RATNLGSLSTAVYNNLRNVLNPKAS-SLITKTLPLNGG-----TLDSNE 222  
 Db 904 SLSTSTGISTSSDNVSKSVSTTERTSVTTITTEPNSSTVSLTOLSLPSVSESS 963  
 QY 223 ITTVNRNNTLTSTINEOKTNAD-----ALNSFIKKVIONNQSQSVFGFTNANVOPSNY 277  
 Db 964 KVFSTNGDQSGSTHQSQSTSTEIEIVTTSKVLPPVSSN-TDLTSEPTNREQPTL 1022  
 QY 278 SFV--AFSADVTVNKKVARTVWNGDEBPSRLIANTNSTIDVSWIYSLAGINTKQVSEF 335  
 Db 1023 STTSNITEDITL-----SQTGNDNDNTSS-----TNPVPIVA-TSTLASASEEDNKG 1071  
 QY 336 SNYGPSTGYLPFYKLVKADANNVGLQYKLNNGVQVEFATSTSNANTNP-----P 391  
 Db 1072 SHESASTS-----LKPSMGNSGL-----TTSTIEATTSPTREAPSP 1109  
 QY 392 ANDEIVAKIVLSGLRFGQNTIELSVPTBGNKKVAPMIGNIYSSNNNADKIPGYR 451  
 Db 1110 AVSSGTDVTEPTDREOPTLS---TTSKTNSESAVATL-----QATNEN-----GGKS 1155  
 QY 452 PGTF 456  
 Db 1156 PSTDL 1160

RESULT 10  
 FNBA STAAU  
 ID FNBA STAAU STANDARD; PRT; 1018 AA.  
 AC P14738;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibronectin-binding protein precursor (FNBP).  
 FNBA.  
 OS Staphylococcus aureus.  
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OK NCBI\_Taxid=1280;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCCTC 8325-4;  
 RX MEDLINE=8909898; PubMed=2521391;  
 RA Signaes C., Raucet G., Joensson K., Lindgren P.-E.,  
 RA Anantharamiah G.M., Hoeseck M., Lindberg M.,  
 RT "Nucleotide sequence of the gene for a fibronectin-binding protein  
 RT from Staphylococcus aureus: use of this peptide sequence in the  
 RT synthesis of biologically active peptides".  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).  
 CC -I- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN  
 CC PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE  
 CC WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO  
 CC THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S. AUREUS,  
 CC THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE  
 CC INVASION.  
 CC -I- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J04151; AAA26632.1; -  
 DR InterPro; IPR004237; Fn\_bind.  
 DR InterPro; IPR005877; Gpos\_Y5IRK.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR006192; LpXTG.  
 DR Pfam; PF02986; Fn\_bind; 1.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; Y5IRK\_signal; 1.  
 DR TIGRfams; TIGR01167; LpXTG\_anchor; 1.

DR TIGRfams; TIGR01168; Y5IRK\_signal; 1.  
 DR PROSITE; P850847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall; Peptidoglycan-anchor; Repeat; Signal.  
 FT SIGNAL 1 36  
 FT CHAIN 37 985 FIBRONECTIN-BINDING PROTEIN.  
 FT PROPEP 986 1018 REMOVED BY SORTASE (POTENTIAL).  
 FT REPEAT 545 574 B-1.  
 FT REPEAT 575 604 B-2.  
 FT DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS,  
 FT REPEAT 745 782 FIBRONECTIN-BINDING DOMAIN.  
 FT REPEAT 782 820 D-1.  
 FT REPEAT 821 859 D-2.  
 FT REPEAT 860 878 D-3.  
 FT DOMAIN 879 948 D-4 (INCOMPLETE).  
 FT REPEAT 879 892 5 X TANDEM REPEATS, PRO-RICH (WR).  
 FT REPEAT 893 906 WR1.  
 FT REPEAT 907 920 WR2.  
 FT REPEAT 921 934 WR3.  
 FT REPEAT 935 948 WR4.  
 FT SITE 982 986 WR5. SORTING SIGNAL (POTENTIAL).  
 FT MOD\_RES 985 985 LPEXTG. LINKED TO CELL WALL (POTENTIAL).  
 SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 5.9%; Score 137.5; DB 1; Length 1018;  
 Best Local Similarity 20.1%; Pred. No. 2.8;  
 Matches 88; Conservative 66; Mismatches 173; Indels 111; Gaps 19;

QY 19 TNSSPSTQNTVSRVSVSSVQSEESTFYLCPPVGSVTVIRLEFGCMSITKQANP--N 75  
 Db 81 TEGSNMTQVTTTEAPVAPVQAPQTAQ-----PANIETVKEE-----VKKEAKPQVKE 128  
 QY 76 NGQTQLEARMELTDLINAKAMTLASLDYAKIFASISASAYSEATVNNINLATELQKM 135  
 Db 129 TTQEQDNSGDQRCYVLDLPKATQNVAVETQVEVAPRTASBKPRTVSADVAEAKASN 188  
 QY 136 AKTNLESAINQANTDKTTFD-----NEHPNLEAVAYAKLTLEQKATNLEGSTAY-- 187  
 Db 189 AKVETGTDV-----TSKYTVEIGSTEGHNNNTKVEPHAGQRAVLKYKLFENGHHQGYFD 244  
 QY 188 ---NQIRNMLVDLYNKA-----SLITKTLPLNGGTLT-----DSNEIT----- 224  
 Db 245 FTLSNNVNTGVSFARKVPEIKNGSVVMATGEVLEGKIKYTFNDIEDKVDYTALEIN 304  
 QY 225 -----TVNENINNTL-STINEOKTN-----ADALNSFIKKVIONNQSQSVFGFTTNA 270  
 Db 305 LFLDPKTVQNTGQITITSTINEBOTSKELDVKKYDGIANYA-----NINGSIETFNKA 358  
 QY 271 NVQPSNYSFVAFSADVTPVNVYKKAARTV-----WNGDEPSRLIANTNSTIDVS-W 320  
 Db 359 N--NRSVAF--IKPNNGKTTSTVVTGLMKGSNONGKQVRILFEYLGNNEDIKAS 412  
 QY 321 IYSLAGINTKYQFSFSNVGPGSTGYLPFYKLVKADANNVGLQYKLNNG-----NVQOVER 376  
 Db 413 VYANTTDTSKFEKVTSM-----SGNLNLQ-----NNGSYSLNIENLDK 451

RESULT 11  
 YKKL CAEEL  
 ID YKKL CAEEL STANDARD; PRT; 1010 AA.  
 AC P34278;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C02F5.1 in chromosome III.  
 GN C02F5.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.

[illegible]

ID	YIG9_YEAST	STANDARD	PRT:	995 AA.
AC	P40442			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical 99.7 kDa protein in SDL1 5' region precursor.			
GN	YIL169C OR Y19402.07C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RA	[1]			
RA	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / AB972;			
RX	PubMed=9169870;			
RA	Churruarín C.M., Bowman S., Badcock K., Bankier A., Brown D.,			
RA	Chillingworth T., Connor R., Devlin K., Gentile S., Hamlin N.,			
RA	Harris D.E., Horsnell T., Hunt S., Jagsels K., Jones M., Lye G.,			
RA	Moule S., Odell C., Pearson D., Rajadream M.A., Rice P., Rowley N.,			
RA	Skellton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;			
RT	"The nucleotide sequence of Saccharomyces cerevisiae chromosome IX."			
RL	Nature 387:84-87(1997).			
CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.ibt-sib.ch/announce/">http://www.ibt-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@ibt-sib.ch">license@ibt-sib.ch</a> ).			
CC	-----			
DR	EMBL; Z46921; CAAB7023.1; -			
DR	PIR; S50358; S50358.			
DR	SGD; S0001431; YIL169C.			
DR	InterPro; IPR004089; Chtax1s transd.			
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2, 1.			
KW	Hypothetical protein; signal.			
FT	SIGNAL 1 23			
FT	CHAIN 24 995			
FT	DOMAIN 26 253			
FT	CARBOHYD 28 28			
FT	CARBOHYD 35 35			
FT	CARBOHYD 468 468			
FT	CARBOHYD 664 664			
FT	SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;			
SC				
Query Match	5.9%; Score 136; DB 1; Length 995;			
Best Local Similarity	17.3%; Pred. No. 3.3;			
Matches	87; Conservative 109; Mismatches 186; Indels 120; Gaps 19.			
QY	18 GTNSS--PSTONTVSRREV-----VSSVQLSEESTFTLCPPGVSTVIRLEFG-----CM 65			
DB	78 GVSSSVGSSSSVSSVSSVSSVSSVSSVSSVSSVSS--SSASDVSSVSSVSSVSSVSS 135			
QY	66 SITKDKANPNNGQQLPEARMELTDLINAKMTLASLDYAKIKIASLSSAYSEAEVTNNN 125			
DB	136 SVSSSSSSASDVSSVSSVSSSSSSASDVSSVSSVSSVSSVSSVSSVSSVSSVSSVSS 188			
QY	126 LNTLEQLKAKTNLESAINQANTDKTTFDNEHNVLVAYKALKTTLEQRATNLEGAST 185			
DB	189 VSSVSSVSSSSASDVSSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSS 248			
QY	186 AYNQIRNNLVLYNKASLLIKTTLDPLNGTILDSNREITVYNNRINNTLSTINEQKTNAD 245			
DB	249 GSGVSS--SSSSSSFPQSTSSASVSSASATSSLSLSTSSASSASATVSSLSSSD 304			
QY	246 -----ALSNSPK-----KVIONNEQSF--VGRTMANVQPSN 276			
DB	305 GTVLPFTTISGDLITGKVLATGEGVVAAGAKTLITLGDGRYSPADIKVGGDLLVKKSK 364			

```

QY 277 YSFVAFSADVTVPVNYKYARTRVWNGDE---PSSRIILNTNSITDVSMTYSLA-GTNTKYQ 332
D 365 ETPGTEPFDISGENPDTVGN--FNAESAAATSAISITSPSSPDNSGDLSLSKSKGE 422
QY 333 FSEFSNTPGSTGYLYFPYKLVKADANNVGLQYKLNNGNVQVFEATST--SANTTANPT 390
D 423 VTFSPYNSGAFSFG-----SNAI-----LNGSGVSGLRDRDTEGSAVNGGIN-- 465
QY 391 PAVDEIKVAKIVLSGLRFGNTIELSVPTGEGNMKVAFMIGNIY----- 436
D 466 -----LDNSTYIVVPPVSGKGVNIIS--GNLYLHYDPTFTGTVVRK 507
QY 437 -----SSENNADKIP--GY 449
D 508 GEGVLAVDPETNATITPVGV 529

RESULT 13
ID MSN2 YEAST STANDARD; PRT; 704 AA.
AC P3378;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein MSN2 (Multicopy suppressor of SNF1 protein 2).
GN MSN2 OR YMR037C OR YMR513.02C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
(1)
(2)
(3)
SEQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Genies S., Hamlin N., Hunt S.,
RA Jagels K., Lyne G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.,
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).
(3)
CHARACTERIZATION.
RP MEDLINE=96208509; PubMed=8641288;
RA Martinez-Pastor M.T., Marchler G., Schueller C., Marchler-Bauer A.,
RA Rius H., Estruch F.;
RT "The Saccharomyces cerevisiae zinc finger proteins Men2p and Men4p
RT are required for transcriptional induction through the stress
RT response element (STRE).";
RL EMBL J. 15:2227-2235(1996).
(1)
FUNCTION: POSITIVE TRANSCRIPTIONAL FACTOR THAT ACTS AS A COMPONENT
OF THE STRESS RESPONSE SYSTEM. RECOGNIZES AND BINDS TO THE
STRESS RESPONSE ELEMENT (STRE) WHICH IS INVOLVED IN THE RESPONSE
TO VARIOUS FORMS OF STRESS (HEAT, OXIDATIVE, OSMOTIC, ETC.).
INVOLVED IN THE REGULATION OF THE CTT1, DDR2, HSP12 GENES.
(1)
SUBCELLULAR LOCATION: Nuclear.
(1)
SIMILARITY: Contains 2 C2H2-type zinc fingers.
(1)
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by, and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/announcements/

```

```

CC or send an email to license@ebi.ac.uk.
CC -----
CC DR EMBL; L08838; AAA34806.1; -
CC DR EMBL; Z48502; CAA88403.1; -
CC DR PIR; S39004; S39004.
CC DR HSPD; P08047; 1SP1.
CC DR TRAF3PAC; T01257; -
CC SGD; S0004640; MSN2.
CC DR GO; GO:0005829; C:cytosol; IDA.
CC DR GO; GO:0005634; C:nucleus; IDA.
CC DR GO; GO:0006950; P:response to stress; IMP.
CC DR InterPro; IPR07087; Znf C2H2.
CC DR Pfam; PF00096; zf-C2H2_2.
CC DR ProDom; PD000003; Znf C2H2_1.
CC DR SMART; SM00355; Znf C2H2_2.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
CC DR DNA-binding; Nuclear protein; Zinc-finger; Metal-binding; Activator;
CC KW Transcription regulation; Repeat.
CC FT DOMAIN 1 109 ASP-RICH (ACIDIC).
CC FT DOMAIN 2 260 279 ASP-RICH (ACIDIC).
CC FT ZN_FING 647 665 C2H2-TYPE 1.
CC FT ZN_FING 676 698 C2H2-TYPE 2.
CC FT ZN_FING 676 698 C2H2-TYPE 2.
CC SQ SEQUENCE 704 AA; 77860 MW; ED6F07446819DF1 CRC64;

Query Match 5.8%; Score 134.5; DB 1; Length 704;
Best Local Similarity 20.0%; Pred. No. 2.6;
Matches 108; Conservative 83; Mismatches 191; Indels 159; Gaps 26;

QY 15 ILVGTNS-----SPSTQ-----NITSREVSVQSEESTPYLCPPVGSVIR 59
D 138 ILVYNNPLOSPPSSSPQNPPTNPINTASNETNLSPTSNGNETL-IPRAQHTSIK 136
QY 60 -----LEFGMSITKQDANPNNGQTOLEARMETDLINAKAMTLASLODYAKIESLSS 114
D 197 DNRSLPNCANSNLIFFITNPNNINEKL--RNQINSPTNSYSNSISN--SNSNSTGLNS 251
QY 115 AYSAEFVNNNNVNTLEQLQAKTNLESAINQANTDKTTD---NEHPVLVAAYALKT 170
D 252 SYFN---SLNDSMDVDVSSDLLNDDDDTNLSRRRSRVITNQFPMTNSRNSIS 307
QY 171 TLE-----QRATNLE-GLSTAYNQIRNNLYDLVYKASSLITKTLDPLNGTILD 219
D 308 SLIDMNHKPINSPNRNTNLTITNTSSSSNASNTITMANNADSNLAG--NPKNDAITD 365
QY 220 SNETTIV---NRNINNTLTSTINEQKTNADALSFTIKKYIQ-----NNEQSF 263
D 366 -NELTOILNEYNNNFNDNLGTSTGK-NKSACPSFDPANMTKINPSQOLQOOLNRFVHK 423
QY 264 VGTFTNANVQPSNYSFAFVSADVTVPVNYKYAR-----TYWNGDE--PSSR 307
D 424 QLTSSHNH--SSTNMKSFNSDL-----YSRQRASLPPIIDSLSDYDLVKNQDEDPKND 474
QY 308 ILANTNSITDVSMTYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVKADANNVGLQYK- 366
D 475 MLFNSNLSGQGIK-----PSMIL-----SDNASVIAKYA 505
QY 367 NNGVQVFEATSTSANNTTANPTPAVDEIKVAKIVLSGL----- 406
D 506 TTSLSDMPELTTEGBON--ANSTPNFD--LSITQNMADLSPASSSSSTSLATNHFYHPP 562
QY 407 RFGQNTIELSVPTGEGNMKVAFMIGNIYLSENN-----ADKIPGYR 450
D 563 QQSHHTMNSKIGSLRRKSAVPLMGTVPILNQNNISSSVNSTGNGAVTKERPSTR 622
QY 451 R 451
D 623 R 623

RESULT 14
PST1_YEAST STANDARD; PRT; 444 AA.
ID PST1_YEAST

```

FT	CARBOHYD	263	263	N-LINKED	(GLCNAC . .)	(POTENTIAL)
FT <td>CARBOHYD</td> <td>268</td> <td>268</td> <td>N-LINKED</td> <td>(GLCNAC . .)</td> <td>(POTENTIAL)</td>	CARBOHYD	268	268	N-LINKED	(GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	280	280	N-LINKED	(GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	292	292	N-LINKED	(GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	305	305	N-LINKED	(GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	329	329	N-LINKED	(GLCNAC . .)	(POTENTIAL)
SO	SEQUENCE	444 AA;	45776 MW;	230F60CACAS521A4	CRC64;	
Query Match						
Best Local Similarity 18.0%; Spec. No. 1.5;						
Matches 89; Conservative 95; Mismatches 185; Indels 126; Gaps 20						
Qy	1 MHYFRNCIFELVILYXGNS--PSQNTSREYVSQVLSSEBETFLCPPEPVSTV	57				
Db	3 LHLSTALTLTSAALATSSSSSSIPSSCTTISHATATA--QSDLKYSRCDTLVNL	59				
Qy	58 I--RLEFGCMSTTKDAPNNQ-----TOLEAPARE-LTDLIAXAMTLASQDY	1058				
Db	60 IGGGLTKGALANVKE-----INGSLTFPAATLTSPADSLSTDSLNQSLTTL	1111				
Qy	106 AKIEASLSAYSEAEFTVNNNLNATLEOLKMAKTLKESAIHQANTDKTTFDNEHPNLEAY	1655				
Db	112 -----SASFGSLQSVDSIKLITLPAISFTSNKSNANNIYISDTSL-----QSVDFG	1588				
Qy	166 KAKTTLLEGQATNLEGLS--TANYQIKMNLVDLYAKRASSLITTLDPPL--NGTGLLDS	2200				
Db	159 SALKKAVFENVVNNKKLTISKSPVETVSDSLQFSFNGNQTKI--TFPDDLWANNISLTDV	216				
Qy	221 NEITTVNRN-----INNTLSTNEOKTVA-----DALSNFIRK-----	254				
Db	217 HASVSPANLOKINSSIGFINNLSINLPIKNTLTIGOTBSIVNDLYKLSNSNLTGGL	276				
Qy	255 VIQNN-----EOSFVGTFTNAVQPSNVSFVAFSADYTPVNYKART	297				
Db	277 VVANNTGLQXIGGLDNLTLTIGTTELVGNFTSLNL--DSLKSVGADVSKSNFSCNA	334				
Qy	298 VNWGDEPSSRIILANTNSITDWSWYISLACTNTKIQFSSNYSPTGYLTPPYLYKRAADA	357				
Db	335 L-----KALQKKGKIGESFVCKNGASSITVLTSTSKQSS-----QTTAKVSKS	380				
Qy	358 NNVTGLQYKLNNGNVQOEPATSTASNTTANPTPAVEIKYAKIVLSGLRFGQNTIELSV	417				
Db	381 SSSAEKKTTSGGITKKAASASVSSGAS-----SSSKSSKGNAAIMA	424				
Qy	418 PTGEGNMNRYAPMIG 432					
Db	425 PIGQ-----TTLVVG 434					
RESULT 15						
Y966 YEAST						
ID	Y966 YEAST	STANDARD;	PRT;	1140 AA.		
AC	004883;					
DT	01-NOV-1997 (Rel. 35, Created)					
DT	01-NOV-1997 (Rel. 35, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DE	Hypothetical 113.1 kDa protein in PRS5-FET4 intergenic region.					
GN	YMR317W OR YMR9924.09.					
OS	Saccharomyces cerevisiae (Baker's yeast).					
CC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;					
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.					
OX	NCBI_TaxID=4932;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=S288C / AB972;					
EX	PubMed=9169872;					
RA	Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,					
RA	Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,					
RA	Rajase K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,					
RA	Ride P., Skelton J., Walsh S., Whitehead S., Barrill B.G.;					
RT	"The nucleotide sequence of Saccharomyces cerevisiae chromosome					
RL	XIII."					
RL	Nature 387:90-93 (1997).					

```

CC -! DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54141; CAA90835.1; -.
DR SGD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match          5.8%; Score 134; DB 1; Length 1140;
Best Local Similarity 21.2%; Pred.No. 5;
Matches 101; Conservative 76; Mismatches 178; Indels 122; Gaps 21;

QY 19 TNSSPSTONVTSREVSVSVQVSEESTFYLCPPVGVSTVIRLEFGCMGSIKKDANPNNGQ 78
DB 412 TSSSVSSSEISSTKSSVMSSEVSATSSLVSSAP--SAISSLASSRLFSK---NTSVTS 466
QY 79 TQLEAARMELTDLINAKAMTLASLDQYAKIEASISAYSEAEYNNNNINATLEQLKMAKT 138
DB 467 TLVATEASSVTSLSLPSSFTLAS---NSIIESSLSTGYN--STVGSTTSAASSTLGSKVS 521
QY 139 NLESAINQANDKTFEDNEHPLVFAKYAKLTLEQRATNL-----EGLSTAYNQ 189
DB 522 SSNSHMAVSKTISTSSDLSKSSVIFGNSVTYTPSPASISLTASPLPSVMSDITSEASS 581
QY 190 IRNNLV-----DLVKKASLITKTLDPPLNGTL--LDSNEITVNRNINNTLSTIN 238
DB 582 ISSNLIASSAPBDNNSTIASASLIYTKXNSVSVSSITSE--TTNESNLATSTSL 640
QY 239 EOKTNADALNSFIKKVIQNEQSF---VGTFTNANVOPSNVSPAFSADVTPVNYKYA 294
DB 641 SNKATARSISTG-----NATSASNVPTGTFP-----SMSHSTSVITP----- 677
QY 295 RRTVNGDEPSSRIIANTNSITDVSMIYSLAGTNTKIQPSFSNYPGPTGYLFPYKLYKA 354
DB 678 -----GFSTSSASLAINSTVVS---SSLAG---YSFSTPESSPPTS-----TLVTS 717
QY 355 ADANNVGLQYKLNNGNVQOVEPATSTSA---NNTTANPTPAVDEIKVAKIVLGLRFGQ 410
DB 718 EAPSTVS-----SMTTSAPFINNSTARPSPSTASF-----ITR 751
QY 411 NTIEL-SVPTGEGNNK-----VAPMIGNIYLSNENNADKI---PGYRRPG 453
DB 752 STSISVPLASGDVTSIAAHNLTFPSAPSTSSAQVLVSKVTSSSILVTPRIDRSG 808

```

Search completed: August 14, 2003, 10:21:39  
Job time : 27 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 14, 2003, 10:18:32 / Search time 101 Seconds

(without alignment)  
1165.069 Million cell updates/sec

Title: US-09-901-572a-3

Perfect score: 2324  
Sequence: 1 MYHRRNCIFFLVILYGTN.....SSNNENADKIFGRYPGFTL 456

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database :

SPTREMBL\_23.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1207	51.9	271	2	Q49464 mycoplasma
2	838	36.1	671	2	Q918D6 mycoplasma
3	828	35.6	702	2	Q49499 mycoplasma
4	825	35.5	632	2	Q9XC68 mycoplasma
5	800	34.4	594	2	Q9KH13 mycoplasma
6	757	32.6	680	2	Q9KH14 mycoplasma
7	750	32.3	702	2	Q49498 mycoplasma
8	746.5	32.1	645	2	Q921D1 mycoplasma
9	745.5	32.1	644	2	Q92HR9 mycoplasma
10	745	32.1	650	2	Q49495 mycoplasma
11	742.5	31.9	486	2	Q8RLX9 mycoplasma
12	741.5	31.9	649	2	Q49497 mycoplasma
13	711.5	30.6	647	2	Q49468 mycoplasma
14	703.5	30.3	703	2	Q05122 mycoplasma
15	703	30.2	656	2	Q9KH15 mycoplasma
16	618.5	26.6	419	2	Q918D5 mycoplasma

17	579	24.9	386	2	Q49500 mycoplasma
18	550.5	23.7	367	2	Q9XC67 mycoplasma
19	314	13.5	805	12	Q98Y44 turkey heip
20	314	13.5	805	12	Q98Y45 turkey heip
21	314	13.5	865	12	Q83291 marek disea
22	314	13.5	865	12	Q83292 turkey heip
23	314	13.5	865	12	Q83293 turkey heip
24	314	13.5	865	12	Q83294 turkey heip
25	195.5	8.4	864	12	Q69408 melegrid h
26	195.5	8.4	864	12	Q69409 melegrid h
27	195.5	8.4	870	12	Q9DP09 melegrid h
28	189.5	8.2	870	12	Q9E1G4 melegrid h
29	181.5	7.8	9439	16	Q49496 mycoplasma
30	177.5	7.6	865	12	Q69406 turkey heip
31	177.5	7.6	865	12	Q9PW21 turkey heip
32	176.5	7.6	933	2	Q53653 staphylococ
33	174	7.5	946	16	Q8NXJ1 staphylococ
34	170.5	7.3	4688	16	Q9PQ08 ureaplasma
35	170	7.3	1302	2	Q49547 mycoplasma
36	170	7.3	6713	16	Q99U54 staphylococ
37	170	7.3	6713	16	Q931R6 staphylococ
38	168.5	7.2	9904	16	Q8NM06 staphylococ
39	166.5	7.2	3890	16	Q99U53 staphylococ
40	166	7.1	661	16	Q8YMD8 anabaena sp
41	164.5	7.1	2481	16	Q99QR6 staphylococ
42	162.5	7.0	810	5	Q814Y6 plasmodium
43	162.5	7.0	1237	16	Q9ZK57 helicobacte
44	162.5	7.0	1461	16	Q8ZL64 salmonella
45	161.5	6.9	1774	5	Q81A08 plasmodium

#### ALIGNMENTS

##### RESULT 1

ID Q49464 PRELIMINARY; PRT; 271 AA.  
AC Q49464;  
DT 01-NOV-1996 (TRENBLREL. 01, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)  
DE TM-1 (Fragment).  
GN TM-1.  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
CX NCBI\_TaxID=2096;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94025893; PubMed=8212828;  
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,  
RA Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.;  
RT "Cloning and DNA sequence of a 23 kilodalton polypeptide gene of  
RT Mycoplasma gallisepticum as a possible protective antigen."  
RL Vaccine 11:1061-1066 (1993).  
DR EMBL; S65869; AAB28343.2; -.  
FT NON TER 271  
SQ SEQUENCE 271 AA; 29817 MW; 8B25DE0CD5C85CA2 CRC64;

QY	64	CMSTTKKDANPNNGOTLEAARMELTDLINKAMTILASLDYAKIEBASLSAYSEAEVFN	123
QY	27	CMSTTKKDANPNNGOTLEAARMELTDLINKAMTILASLDYAKIEBASLSAYSEAEVFN	86
QY	124	NNLVATLEQLMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGS	183
QY	87	NNLVATLEQLMAKTNLESAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGS	146
QY	184	STAYNOIRNNIVLDYKNAASSLITKTLDPPLNGTILDSNEITTVNRNINNTISTINEQKTN	243
QY	147	STAYNOIRNNIVLDYKNAASSLITKTLDPPLNGTILDSNEITTVNRNINNTISTINEQKTN	206

OY 244 ADALNSFIKVIYONNEOSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTVWNGDE 303  
 DB 207 ADALNSFIKVIYONNEOSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTVWNGDE 266  
 OY 304 PSSRI 308  
 DB 267 PSSRI 271  
 RESULT 2  
 Q9L8D6 PRELIMINARY; PRT; 671 AA.  
 AC Q9L8D6; TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE PMGA-like protein 9.2.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NC NCB1\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F,  
 RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,  
 RT May J.D., Hugliett M.B.,  
 RT "A novel PMGA-like gene from the F-strain (vaccine strain) of  
 RT Mycoplasma gallisepticum."  
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF210770; AAF29524.1; -  
 SQ SEQUENCE 671 AA; 71898 MW; 4406A6B591750679 CRC64;

Query Match 36.1%; Score 838; DB 2; Length 671;

Best Local Similarity 40.7%; Pred. No. 2.3e-32; Matches 184; Conservative 78; Mismatches 128; Indels 62; Gaps 10;

OY 53 VGSYTRLEFGCMSTTKDANP-----NNGOT-----OLEEA 84  
 DB 15 ISSFWLAAASCTSATTPPNEPEKDPMPNPPSGDMNGDTPNPGDGMENSAOQLAA 74  
 OY 85 RMELTDLINAKMTLASLQDYAKIEASLSAYSEAEVNNNNLNATLEOLKMAKTNLESAT 144  
 DB 75 KKEISDLATQSNLAKADYNTNIGTTLTAAYTTAKSTSDNTSVLEQVKSATSTLQAAI 134  
 OY 145 NOANTDKTTFDNEHNVLEAYVALKTTLEQKATNLEGLSTAYNOIRNVLDLYNKASL 204  
 DB 135 DPAASKTSFDEKNPELIKAYVALKELTKEEIVLSGLTDSNFATIKTNLTALYOSGDF 194  
 OY 205 IKTLDLNGGTLTDSNEITTVNRNINNTLSTINEOKTNADALNSFIKVIYONNEOSFV 264  
 DB 195 VKATLDPVSGNA-PQIADITKADKDIADAVSKLETWKTNANTLATSFEVAVKUKLIGI 253  
 OY 265 GFTFNANVOPSNYSFVAFSADVTVP---VNYKARRTVWNGDEPSSRIIA---NTNSIT 316  
 DB 254 DT-TNNREPGNVYSFVGYVNAATNNEIPIWNAFAQKRWTSNDGRTSLISTSDMSITLT 312  
 OY 317 DVSMTISLAGTNTKYOPSSSNVSPSGYLYFPYKLVKADANNVGYOYLNGNNGVQVAF 376  
 DB 313 EVSWIYSLSGAGTKYSLTFNYGPGSTGYLFPYKLVKEGENNVVALQYTLNNGSAQOVNF 372  
 OY 377 A-----TSTSA-----NNTT-----ANPTPAVDEIKVAKIVLSGLRFGONTLE 414  
 DB 373 APVTKTSVADSGDSDNNQTESAAETMPVTSIDNPATVSDINIAITLISNLFSGSTIE 432  
 OY 415 LSVPTGEGNNKVAAPMIGNTYSSNENNDKI 446  
 DB 433 FSVPTEPS--NKVAPMIGMYLTSNLANEAKV 462

RESULT 3  
 Q49499 PRELIMINARY; PRT; 702 AA.  
 ID Q49499;  
 NC Q49499;

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PMGA1.4 protein precursor.  
 GN PMGA1.4.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NC NCB1\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S6;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
 RA Browning G.F., Whitehead K.G., Walker I.D.,  
 RT "The organisation of the multigene family which encodes the major cell  
 RT surface protein, PMGA, of Mycoplasma gallisepticum."  
 RL FEBS Lett. 352:347-352(1994).  
 DR EMBL: L28424; AAA62418.1; -  
 KW Signal.  
 FT SIGNAL  
 SQ SEQUENCE 1 25 POTENTIAL.  
 B70AC874FE85055C CRC64;

Query Match 35.6%; Score 828; DB 2; Length 702;

Best Local Similarity 40.6%; Pred. No. 7.3e-32; Matches 176; Conservative 86; Mismatches 125; Indels 46; Gaps 9;

OY 50 PPPGVSIVILFEGCMSTTKDANPNNG-----QTOLEAARMELTDLINAKMTLASLQ 103  
 DB 54 PNPFGGMMGGNGG-----NTNPGNGGGTDNAAQQLAAKKEISDLATQSNLSTYA 106  
 OY 104 DYAKIEASLSAYSEAEVNNNNLNATLEOLKMAKTNLESAINOANTDKTTFDNEHNVLE 163  
 DB 107 DYANIGTTLTAAYTTAKSTSDNTSATLEQVKSATSTLQIADTAAASKTSFDEKNPELIK 166  
 OY 164 AYVALKTTLEQKATNLEGLSTAYNOIRNVLDLYNKASLITETLDPINGTLTDSNEI 223  
 DB 167 AYVALKTTLEQKATNLEGLSTAYNOIRNVLDLYNKASLITETLDPINGTLTDSNEI 225  
 OY 224 TTVNRNINNTLSTINEOKTNADALNSFIKVIYONNEOSFVGTFTNANVOPSNYSFVAFS 283  
 DB 226 SQANTNINNAVSKLETWKTNANTLATSFEVAVKUKLIGIT-TNNOEGPGNVSPGYVS 284  
 OY 284 ADVTP-----VNYKARRTVWNGD-----EPSSRIILANTNSITDVSMTISLAGTNTKY 332  
 DB 285 VDVTTGSDNAPPNMSPAFQKRWTSNTDILSQPQAEGENQOSADVSIVYVLTGAKYYS 344  
 OY 333 FSPSNVSGSTGYLYFPYKLVKADANNVGYOYLNGNNGVQVAFS-----TSAN 383  
 DB 345 LTFNYGPGSTGLFPYKLVNVSDDKVALLEYKUNESAANKTIDSPSQTSPVADATREN 404  
 OY 384 N-----TTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAAPMIG 432  
 DB 405 NRSTAPAPQSGSTEINPATLTDIKAKVTLNLFKFSNTIEFSVPTTAKEGTSKVAAPMIG 464  
 OY 433 NIYSSNENNDK 445  
 DB 465 NMVLTSSDRDYVK 477

RESULT 4  
 Q9XC68 PRELIMINARY; PRT; 632 AA.  
 ID Q9XC68;  
 AC Q9XC68; TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE VHA1 precursor (Fragment).  
 GN VHA1.  
 OS Mycoplasma imitans.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NC NCB1\_TaxID=29560;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=4229; PubMed=10463176;  
 RA MEDLINE=99392472; Duffy M.F., Glew M.D., Browning G.F.;  
 RT "A gene family in Mycoplasma imitans closely related to the pmcA  
 RL family of Mycoplasma gallisepticum";  
 RU Microbiology 145:2095-2103(1999).  
 DR EMBL, AF141940; AAD39463.1; -.  
 KW Signal.  
 FT NON TER 1 17 POTENTIAL.  
 FT SIGNAL 18 632 VLHAI  
 FT CHAIN 18 632  
 SQ SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;  
 Query Match 35.5%; Score 825; DB 2; Length 632;  
 Best Local Similarity 43.4%; Pred. No. 8.9e-32;  
 Matches 187; Conservative 68; Mismatches 130; Indels 46; Gaps 11;

QY 53 VGSIVIRLEPGC-MSITKCDANPNN-----GQT---QLEAARMELTDLIN 93  
 Db 7 IGSFVTLAAASCSKAVTSPNCPNPNPKFNNPDGTGNTPGGDTDAKQIADAKASLNTLLG 66  
 QY 94 AKAMTLASLODYAKIEASISAYSEAEFTVNNNLNATLEQLKMAKTNLESAINQANTDKT 153  
 Db 67 SQSTNVVALYEDYAKIKOTLSAIVASAOQTANNANATLKVDADAKAQTALSDANAKTE 126  
 QY 154 PDNEHPNLVEAYKALKTLEQATNLEGLSSTAYNOIRNNLVLYNKASSLITKLDPLN 213  
 Db 127 FDKANLGLVSAVAKETLKESTNTDLSQDNYSALKANVSLYNKAKDTTYTLDL-T 185  
 QY 214 GGTLLDSNEITTVNINNTLSTINEOKTNADALSNSFIKVIQNNESFVGTFTN---- 269  
 Db 186 SGMIPIKDEITSANTALITQAVSAIDSOCTNADTAATFEIKEKIDSAKLT-EGTAAEGQA 244  
 QY 270 ANVQSNVSPVAFPSADYTP-----VYKKYARRYVNNNGDEPS--SRLANTNSIT 316  
 Db 245 ASQPGNSFVGFSDVDTGTGTGSOEDLPFNMFAKRVMTTEGLSACQTLVSESTPLT 304  
 QY 317 DVSMTYSLAGTNTKYQSFNSNYGPGSTGYLYFPYKLVKADANNVGLQYKLNNG-NVOQVE 375  
 Db 305 DVSMTYSLTGAGSKYTLFTFYGPSTALVYFPYKLVQSSDNKGLQYKLNNSSETIVPIT 364  
 QY 376 PATSANTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAEMIGNIY 435  
 Db 365 FGNFTN----DSGATPAIDIKVEKVTLSNINFGDNTIEFSVATEE--NKVAPMIGMY 417  
 QY 436 LSSNENADKI 446  
 Db 418 LTSSTNNVDKI 428

RESULT 5  
 Q9KH13 PRELIMINARY; PRT; 584 AA.  
 ID Q9KH13  
 AC Q9KH13  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 RT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Adhesin pmcA1.4 (Fragment).  
 GN pmcA1.4.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HS;  
 RA Shen Q.C., Bi D.R., Weng C.J.;  
 RT "Sequence analysis of the pmcA multigene family of Mycoplasma  
 RL gallisepticum strain HS";  
 DR Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF275312; AAF91415.1; -.  
 DR InterPro: IPR001986; EPPS\_synase.  
 DR PROSITE, PS00104; EPPS\_SYNTHASE\_1; 1.  
 FT NON\_TER 584 584

SQ SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;  
 Query Match 34.4%; Score 800; DB 2; Length 584;  
 Best Local Similarity 42.1%; Pred. No. 1.3e-30;  
 Matches 179; Conservative 71; Mismatches 115; Indels 60; Gaps 10;

QY 73 NPNNGQT---QLEAARMELTDLINAKAMTLASLODYAKIEASISAYSEAEFTVNNNLN 128  
 Db 69 NPNNGTTPBQQLAAAKKTLTDLGENTNVVALYADYAKIQSTLSTAYMTAKTASENTSA 128  
 QY 129 TLEQAKAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTLEQATNLEGLSSTAYN 188  
 Db 129 TLEQAKAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTLEQATNLEGLSSTAYN 188  
 QY 189 QIRNNLVDLYNKASSLITKLDPLNGTLLDSNEITTVNINNTLSTINEOKTNADALS 248  
 Db 189 SIKHLSKLPDASALAKTLDP-TMGITVEWNSVTANEDIMTAVSKLTETKTNADKF- 246  
 QY 249 NSFPIKVIQNNESFVGTFTNAN--VQSNVSPVAFPSADYTP-----VYKKYART 297  
 Db 247 NDFEKKPL--SKELVSTNDRAHNQEPANWSPAGYVDLTGSGTNSQNLPMNPFQPK 304  
 QY 298 VVNGD---EPPSRLANTNSITDYSMTYSLAGTNTKYQSFNSNYGPGSTGYLYFPYKLVKA 354  
 Db 305 VMTSEGOQTGKTALVSSPVASATDVSWIYSLAGETKYTLFFEYYPDPJALYFPYKLVKQ 364  
 QY 355 ADANNVGLQYKLNNGVQOVEF-----ATST-----S 381  
 Db 365 ADSSVALQYSLNKTSSKLINFPAKTMPTNADSENGVATTSTTEGRSSSEVLVADEVA 424  
 QY 382 ANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAEMIGNIYLSNEN 441  
 Db 425 AVNNENMPTPLVSINIAKVTLSGLTGEINTIEFSVPT-----NKVAPMIGMYLTSNSG 479  
 QY 442 NADKI 446  
 Db 480 SGGKI 484

RESULT 6  
 Q9KH14 PRELIMINARY; PRT; 680 AA.  
 ID Q9KH14  
 AC Q9KH14  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 RT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE Adhesin pmcA1.3.  
 GN pmcA1.3.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HS;  
 RA Shen Q.C., Bi D.R., Weng C.J.;  
 RT "Sequence analysis of the pmcA multigene family of Mycoplasma  
 RL gallisepticum strain HS";  
 DR Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF275312; AAF91414.1; -.  
 DR EMBL, AF275312; AAF91414.1; -.  
 SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;  
 Query Match 32.6%; Score 757; DB 2; Length 680;  
 Best Local Similarity 39.6%; Pred. No. 1.7e-28;  
 Matches 163; Conservative 71; Mismatches 130; Indels 48; Gaps 7;

QY 71 DANPNNGQTOLEAARMELTDLINAKAMTLASLODYAKIEASISAYSEAEFTVN 123  
 Db 63 DTVNGGQGMNATNGLVNAKKALSDLIGESKYVELYADYAKIKADLTSAYVAKTTS 122  
 QY 124 NNANATLEQLKMAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTLEQATNLEGLS 183  
 Db 123 DSSTSTLDQVKTATSTLTQTAINTAASDKEPQDSQSLAAVYKLVKOTLNKKEAIVMSLN 182

```

QY 184 STAYNOIRNNLVLDYNNKASLITKTLDPNGTLLDSNEITTTNRNINNTL-----ST 236
DB 183 QEKYSALISEINAASTAEIIVKOTLNVNG-----NLPVAAALMENTKILEAIKEEK 236
QY 237 INEKTADAASNFPIKVIQNNQSPFGFTNANVOPSNYPFAFSAADVPVNYKYARR 296
DB 237 INSEKSNADLFPANYQYKTL--DRTKLMSEGSNNTKQPGNYSFAYASDIASPRMNAQR 293
QY 297 TVMNGD-----EPSSRIANTNSITDVSIMYSLAGTNTKYOFSNNGPSTGYLPPYKL 351
DB 294 TVMTADSRWTSPLEPNLQNGAPLTDVSMYTLSTGAKYTLTFDYGPQGTGYLPPYKL 353
QY 352 VKAADANNVGYOKLNNNGVQVEF-----ATSSANTTANPTTNPAY 393
DB 354 VKTSD--KVGLOQYKLNQADPVALIOFSEAAATASAPAEITDGOESAETATANEKVPMSV 411
QY 394 DEIKVAKIVLSGRFGONTIELSVPTGNNKRYAPMIGNIYILSSNENNAKD 445
DB 412 NTINVAKVTLSNLRKGSNTIEFSVPMDDNNKTAIPMIGNIYITSSNDEANK 463

```

## RESULT 7

```

Q49498 PRELIMINARY; PRT; 702 AA.
ID Q49498:
AC Q49498:
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DE PMGAL.3 protein precursor.
GN PMGAL.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6;
RA MEDLINE=95010739; PubMed=7925999;
RA Marham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whitehead K.G., Walker I.D.,
RA "The organisation of the multigene family which encodes the major cell
RA surface protein, PMGA, of Mycoplasma gallisepticum.";
RT FEBS Lett. 352:347-352(1994).
RL EMBL: L28424; AAA62417.1; -.
KM Signal.
FT SIGNAL.
SQ SEQUENCE 702 AA; 75537 MW; 273E8915FE57B9F CRC64;

```

Query Match 32.3%; Score 750; DB 2; Length 702;  
 Best Local Similarity 41.4%; Pred. No. 3.7e-28;  
 Matches 175; Conservative 62; Mismatches 128; Indels 58; Gaps 11;

```

QY 73 NENNGQT---OLEAARMELTDLINAKAMTASLQDYAKIEASLSANSEATTNNNINA 128
DB 69 NENPGNTPEQOLAARKTLTDLTGENTNNVALYDIAKIOSTLTAVYATATAEINTSA 128
QY 129 TLEQKMAKTNEESAINQANTDKTFPNDNEHPLVENVYKALKTLLEQATNLEGLSTAYN 188
DB 129 TLENRASATTLQALIDKANDKRVFDSNOPLVAAYNNKLTLLSKTISLGLSENKYG 188
QY 189 QIRNNLVLDYNNKASLITKTLDPNG--GTLDSNEITTVNNINNTLS--TINEOKTVA 244
DB 189 GIKNHLSKLPDGSATLAKTLDPTSGERTLEKVE--ANNGIKMAISPELSKWKXNA 245
QY 245 DALNSFIKVIQNNQSPFGFTNANVOPSNYPFAFSAADVP-----VNYKYARRTW 299
DB 246 DLF--NEFEKRPISKEKLSSTDAHQEQEPANWSFPAVSVDLTSSQNLPPNNPQORKW 304
QY 300 NCD--EPPSRILANTN--SITDYSWISLAGTNTKYOFSNNGPSTGYLPPYKLVKAD 356
DB 305 TSENQPGKTAIVSSVSATIDVSMYSLAGEGTKYTLTFEYGPDAAPFLYLYKLVKAD 364
QY 357 ANNNGLOQYKLNNGVQVEF-----ATST-----SAN 383

```

```

DB 365 SSSVALQYSLNKTSSKLNPKPAETVSTNTDSENEVATSTTEARSSYKVLVADEAAT 424
QY 384 NTTANPTPAVDEIKVAKIYLSGRFGONTIELSVPTGEGNNKRYAPMIGNIYILSSNENNA 443
DB 425 NNMENHTPTVDINIAKVTLISGLTGENTIEFSVPEG-----KVAPMIGNIYILSSNESQ 479
QY 444 DK1 446
DB 480 VK1 482

```

## RESULT 8

```

Q92ZDI PRELIMINARY; PRT; 645 AA.
ID Q92ZDI:
AC Q92ZDI:
DT 01-MAY-1999 (TRENBLREL. 10, Created)
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
DE M9 protein.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31;
RA MEDLINE=9900182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.,
RA "A protein (M9) associated with monoclonal antibody-mediated
RA agglutination of Mycoplasma gallisepticum is a member of the PMGA
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;

```

Query Match 32.1%; Score 746.5; DB 2; Length 645;  
 Best Local Similarity 40.4%; Pred. No. 4.9e-28;  
 Matches 176; Conservative 67; Mismatches 134; Indels 59; Gaps 12;

```

QY 53 VGSIVTILLEGCMSTK-----DANPNNGQ-----TOLEAARMELTDLINA 94
DB 15 IGSFVMLAASCTPTPNPNPPSGMNGGDTNPEDGGMMNAASQELAAARMGLTTFDS 74
QY 95 KAMTASLQDYAKIEASLSAYSEATTNNNLANLTLEQKMAKTNEESAINQANTDKTF 154
DB 75 KAKVLGYVDYKKTQDTLTAAYDAKTYVLDNSSSTONLINEAKTRLETAIRTAATSKOTF 134
QY 155 DNEHPNIVEAYKALKTLLEQATNLEGLSSTAYNQIRNNLVLDYNNKASLITKTLDPNG 214
DB 135 DEQHAELVYKVEELKTLTSMETATLAPYAAQVAGIKHLSGLYDAGKALITKLEPVEG 194
QY 215 GTLDSNEITTVNNINNTL--STINEOKTVAADLSNFIKVIQNNQSPFGFTNANV 272
DB 195 DP--LTADVMMANKIYEAIDDEVLANPOKENATLDAISFVQYVYKKTIVEEAHN--KA 252
QY 273 QPSNYPFAFSAADVP-----NRYKARRTW--NGDEPSSILANT-----NSI 315
DB 253 QPANYSFVGYVDITGVGTQTSIPNDYARITFTNGDEP--RSISNTPADGGTMOPL 310
QY 316 TDVSWISLAGTNTKYOFSNNGPSTGYLPPYKLVKAAADANNVGYOKLNNNGVQVE 375
DB 311 SNVSWISLAGTNGAKYTLTFEYGPSTGYLPPYKLVNTSOMKGLGLEYKLAND----- 363
QY 376 FATSTSA-----NNTTANPTPAVDEIKVAKIYLSGRFGONTIELSVPTGEGNNKRYAPM 430
DB 364 -ATEPSALTTPSEQTMNCKTPTVDINIAKVTLANLNRKGSKIEFSVPA-----EKVSPM 417
QY 431 IGNITLSSNENNAKD1 446
DB 418 IGNITLSSSPNNMNKI 433

```

## RESULT 9

Query Match	32.1%;	Score 745.5;	DB 2;	Length 644;
Best Local Similarity	40.4%;	Pred. No. 5.5e-28;		
Matches 176;	Conservative 68;	Mismatches 133;	Indels 59;	Gaps 12

RESULT 10  
Q49495  
ID Q49495 PRELIMINARY; PRT; 650 AA

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Haemagglutinin precursor.  
GN MGCA1.1  
OS Mycoplasma gallisepticum.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma  
OX NCBI\_TaxID=2096;  
RN [1]

Query Match	32.1%;	Score 745;	DB 2;	Length 650;
Best Local Similarity	40.5%;	Pred. No. 5.9e-28;		
Matches 177;	Conservative 65;	Mismatches 139;	Indels 56;	Gaps 12;

RESULT 11  
Q8RLX9  
ID Q8RLX9 PRELIMINARY; PRT; 486 AA

GN VtLA.. *Mycoplasma gallisepticum*.  
OS *Bacteria*; *Firmicutes*; *Mollicutes*; *Mycoplasmataceae*; *Mycoplasma*  
OC NCBI\_TaxID=2096;  
OX [1]  
RN RP SEQUENCE FROM N.A.  
RC STRAIN=tc-11;  
RA Matham P.F., Kanci A., Zaïra G., Sundquist B., Haines P.,  
RA Browning G.F.,

"Malt homolog in Mycoplasma gallisepticum is not essential in vitro."

RT Submitted (JDC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY055985; AAL58980.1; -

FT NON\_TER 486 486

SO SEQUENCE 486 AA; 51844 MW; 94912DD7A09FB911 CRC64;

Query Match 31.9% Score 742.5; DB 2; Length 486;

Best Local Similarity 38.0%; Pred. No. 5.5e-28;

Matches 175; Conservative 68; Mismatches 128; Indels 89; Gaps 12;

51 PVGVTIRLEFGCMSTTKK-----DTQLEAARMELTDLNAKAMTLASIOD 104  
 52 PPSGG-----NNMGSDTNPENGGGMNANOVLAAKATLTLLNGTEKVGLYND 101  
 105 YAKIETASISAYSEAEYNNNNLNATLEOLKAKNTLESAINQANTDKTTPNEHPNIVEA 164  
 102 YAKIETDLYKAYIAKEISDSKSHATLOEVNNAKRTLETAKDAANSKTSFEEKPELIKKA 161  
 165 YKAKTTLLEQRATNLEGLSTRAYNOIRNNVLDLYNKASLITKTLDPINGTLDSNBIT 224  
 162 YDALKQITISEBMSINQIMDANFETIKHISNLYKQGDIIITATLDPPTG---DGPQAM 217  
 225 TYNRN---INNTLTINQKTNADALNSFIKGVION---NEQSFVGTFTNANVQPSNY 277  
 218 VNQGTNEALVNAKTSKIEDKTNATLAFVQKTLNANLVNE-----TNNQPSQSY 270  
 278 SFVAESADV-----TPVNYKYARFTV---NGDEP---SSRIANTNSITDVSVITY 322  
 271 SFVAVSVLDNTGVSTASNTF-NMNLAQKRVWVSGSGRTSPSSSDANNSPALTDVSWIT 329  
 323 SLAAGNTTKQSFSPSYNGPSTGYLYFPYKLYKAAADANNVGLQKLNNGVQVEFA----- 377  
 330 NLSGANSKTYLTFNMYGPGSTGHLFPPYKLVKSDQONVGLQTLTNKPKAQRIEFPAPOSP 389  
 378 -----TSTSNNTTANPTPAVDEIKVAKIVISGL 406  
 390 SSGGTAHADPQSPRAAATETVDSASBSQAOTIMSSSMNKKTFTVDINAVATLSDL 449  
 407 RFGQNTIELSVPTGEGNNKRVAPMIGNIYLSNENNAKDI 446  
 450 NFGANTIEFSVPMGD---SMVAPMIGNMYITSNPLNVOI 486

# RESULT 12

Q49497 PRELIMINARY; PRT; 649 AA.

AC Q49497; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE PMGAL.2 protein precursor.

GN PMGAL.2

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S6;

RX MEDLINE=95010739; PubMed=7925999;

RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,

Browning G.F., Whitehead K.G., Walker I.D.;

"The organization of the pmga, of Mycoplasma gallisepticum,"

RT surface protein, pmga, of Mycoplasma gallisepticum,"

RL FBS Letc. 352:347-352(1994).

DR EMBL; L28424; AAA62416.1; -

KM Signal.

FT SIGNAL 1 25

SO SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 31.9% Score 741.5; DB 2; Length 649;

Best Local Similarity 40.0%; Pred. No. 8.6e-28;

Matches 176; Conservative 65; Mismatches 136; Indels 63; Gaps 12;

53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90  
 15 IGSFVWLAASCTTPTNPPTNPSPSGGMNGDTNPGQGGMMNAASQELAAARVGLTT 74  
 91 LINAAMTLASIODYAKIEBASISAYSEAEYNNNNLNATLEOLKAKNTLESAINQANTD 150  
 75 IPDSKAKNIGLYVDYKKTQNTLTKAYDAKTVLDNSSSTTONINEAKRTLETAFRTAATS 134  
 151 KTFPNEHPNIVEAYKALKTTTLEQRATNLEGLSTRAYNOIRNNVLDLYNKASLITKTLTD 210  
 135 KQTFDQHELVKVEYELKTLTSLNETATLAPYADAOYAGIKHLSLYAGAAITTKTLE 194  
 211 PLNGGTLDNSNETTYNRRNNTL---STINEQKTNADALNSFIKGVIONNEQSFVGTIT 268  
 195 PVGSDP-LTASAMAMANTKIVEAKIOEVLNPGKENTKLADSPVKQVLKAKITGEBAH 253  
 269 NANVQPSNYSFVAFSADVTPV-----NYKARFTV-NGDEPSSRIANT- 312  
 254 N-KAQPANYSFVGSVDITGTTGQTSIPNWDYAQRTITNSDEP-RSISNTPADGQTM 310  
 313 -NSITVSWITSLAGNTTKQSFSPSYNGPSTGYLYFPYKLYKAAADANNVGLQKLNNGV 371  
 311 AQPUSNVSWITSLAGNTTKQSFSPSYNGPSTGYLYFPYKLYKAAADANNVGLQKLNNGV 367  
 372 QOVEPATSTSA-----NNTTANPTPAVDEIKVAKIVISGLRFGQNTIELSVPTGEGNNK 426  
 368 -----ATKPSAITFGSDQTNNGKTPVNDINAVKVLNANLFGSNKIEFSVPA-----EK 417  
 427 VAPMIGNIYLSNENNAKDI 446  
 418 VSPMIGNMYITSNPLNVOI 437

# RESULT 13

Q49468 PRELIMINARY; PRT; 647 AA.

AC Q49468; 053303;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Hemagglutinin homolog precursor.

GN PMGAL.2

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=9162830; PubMed=8432610;

RA Markham P.F., Glew M.D., Whitehead K.G., Walker I.D.;

"Molecular cloning of a member of the gene family that encodes pmga, a

hemagglutinin of Mycoplasma gallisepticum,"

RL Infect. Immun. 61:903-909(1993).

DR EMBL; M83178; AAA02996.1; -

DR EMBL; S55216; AAB25397.2; -

KW Signal.

FT SIGNAL 1 25

FT CHAIN 26 647

SO SEQUENCE 647 AA; 70333 MW; 33916673B9528C4 CRC64;

Query Match 30.6% Score 711.5; DB 2; Length 647;

Best Local Similarity 39.0%; Pred. No. 2.3e-26;

Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;

53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90  
 15 IGSFVWLAASCTTPTNPPTNPSPSGGMNGDTNPGQGGMMNAASQELAAARVGLTT 74  
 91 LINAAMTLASIODYAKIEBASISAYSEAEYNNNNLNATLEOLKAKNTLESAINQANTD 150  
 75 IPDSKAKNIGLYVDYKKTQNTLTKAYDAKTVLDNSSSTTONINEAKRTLETAFRTAATS 134  
 151 KTFPNEHPNIVEAYKALKTTTLEQRATNLEGLSTRAYNOIRNNVLDLYNKASLITKTLTD 210

Db 135 KOTPEOHAELVKYKELKTLTSLNETATLAPADAOVAGIKMHLGSLYDACKAITTKLE 194  
 Qy 211 PLNGGTLDSNEITVNNINNTL--STINEOKTN---ADALSNFICKYIIONNEQSFV 264  
 Db 195 PVEBDP-LTASAVMANAKIYEALDEVLNPOKENTATLADSLSLYKITTGVEE---- 249  
 Qy 265 GTFTNANVPNSYFVAFSADVTVP-----NYKARFTW-NGDEPSSRIANT--- 312  
 Db 250 ---AHNKQAPANSYFVGKRWYTELTLDDKQVEPNMVDYQRTIFTSNDEP--RSISNTPAD 304  
 Qy 313 -----NSTDYSWYSLAGTNTKYQFSNSNGPSTGYLYFPYKLVKADANNVGLQYTLN 367  
 Db 305 GQTAQAPANSYFVAGTAKTLETTETTYGPGTGYLYFPYKLVNTSDQVYKLGLEYTLN 364  
 Qy 368 NGNVQOVEFATSTSA-----NNTTANPTPAVDEIKAKIVLSGLRFGQNTIELSVPTSEG 422  
 Db 365 D-----ATKPSAITFGSDQTNMGKPTFYNDIVAKVTLANLNFSGNKIEFVSIPA--- 413  
 Qy 423 NNNKVAPEMIGNIYLSNENNAKDI 446  
 Db 414 --EKVSPMIGNMYLSSSPNNMKI 435

RESULT 14

005122 PRELIMINARY; PRT; 703 AA.  
 ID 005122;  
 AC 005122;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Haemagglutinin.  
 GN PMGAL.9.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Glew M.D., Sykes J.B., Bowden T.R., Pollocks T.D.,  
 RA Browning G.F., Whithear K.G., Walker I.D.,  
 RT "The organisation of the multigene family which encodes the major cell  
 RT surface protein, pmga, of Mycoplasma gallisepticum.";  
 RL FEBS Lett. 352:347-352(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56;  
 RA Markham P.F.,  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90714; AAB50154.1;  
 SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CBC5 CRC64;

Query Match 30.3%; Score 703.5; DB 2; Length 703;  
 Best Local Similarity 35.0%; Pred. No. 66-26;  
 Matches 171; Conservative 77; Mismatches 125; Indels 115; Gaps 12;

Qy 53 VGSFTVIRLEFGOMSTTKDAMP-----NNGOT-----QLEPAA 84  
 Db 15 IGSFPMALAAAGTSAITITPLNFTNPEBPKPDBMPNPGGANGGNTNGMDTAOELASA 74  
 Qy 85 RMEITLDLINAKAMTLASLDYAKIEASLSAYSEAEVNNINLATELQKAKTNLBSAI 144  
 Db 75 KAALITTLNRSSEKVGLYVDYAKIKADLTSAVTAKTSDSSTSLVQVKATSLQTAI 134  
 Qy 145 NQANTDKTTPDNEHNVLEAYKALKTITEFOR-ATNLEGLSTAYNQIINNVLVDLYNKAS 203  
 Db 135 DKASDQKQFQDHDLDLMPYSELKTTLSQKNAATL--LNOPKXSAIINKNSIYAGEE 192  
 Qy 204 LITKTLPLNGGTLDSNEITVNNINNTLS--TINOKTADALNSFIKVIQONNEQ 261  
 Db 193 VVIRLDPVVS-CAITTAASITKNDKINAKISENOLKPKKONADAPFANYQPFKL---DXT 248  
 Qy 262 SFVGTFTNANVPNSYFVAFSADVTVP-----NYKARFTWNGDEPSSRIANTN 313

Db 249 KIMGSTMNMKQPNQYFVGYSVGTGQWOSQTTIPWNNFAQRIWSSGAPRADLASQTE 308  
 Qy 314 S----- 314  
 Db 309 TPOAETPPMSAPQCVEPAAQOQDSSPKOASFTQEVSPTPPAEVOAQADTEQATSGCTP 368  
 Qy 315 ITDVSMTYSLAGTNTKYQFSNSNGPSTGYLYFPYKLVKADANNVGLQYTLNNGVQV 374  
 Db 369 LTDVSWIYSLSGTIVKKTFTFNNGPSPMAYLYFPYKLVKSD--SVGLQYTLNNNNPYAL 426  
 Qy 375 EPATSTSNANTTANPTPAVDEIKAKIVLSGLRFGQNTIELSVPTSGGNMKVAPMIGNI 434  
 Db 427 NFGSETNAN--GPAASVDNINAKVNLNANGENTIEFVSVP-----NNKVAPEMIGNM 477  
 Qy 435 YLSNENNA 442  
 Db 478 YITSDVAN 485

RESULT 15

09KH15 PRELIMINARY; PRT; 656 AA.  
 ID 09KH15;  
 AC 09KH15;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE Adhesin pmga1.2.  
 GN PMGAL.2.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HS;  
 RA Shen Q.C., Bi D.R., Weng C.J.,  
 RT "Sequence analysis of the pmga multigene family of Mycoplasma  
 RT gallisepticum strain HS."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF25312; AAP91413.1;  
 SQ SEQUENCE 656 AA; 70875 MW; 58EA7B075FCC17B1 CRC64;

Query Match 30.2%; Score 703; DB 2; Length 656;  
 Best Local Similarity 40.7%; Pred. No. 5.9e-26;  
 Matches 166; Conservative 68; Mismatches 120; Indels 54; Gaps 13;

Qy 71 DANPNNGOT-----QLEAPMELTDLINAKAMTLASLDYAKIEASLSAYSEAEVYN 123  
 Db 63 DINGGQNMWDSAAOELTARTRALTSILASKNANVEYSDYAKIQHTLAAVTTAEGTS 122  
 Qy 124 NNINATLEOLKMAKTNLESAINOANTDKTTPDNEHPNULVEAYKALKTTLBOGATNLBGLS 183  
 Db 123 QNSSATLEQVNAATSAIQATALNTANSNKKQFDQDSNLMSYKULMATLAKKETAVMTLK 182  
 Qy 184 STAYNOIRNNLVLDLYNKASSLITKTLPLNG-----GTLDSNEITVNNINNTLSTIN 238  
 Db 183 DPKYSALIDQNGVSSNGBELVQHTLDPVSGIVPAANTI--TEBITKIEVYISE--KTLQ 238  
 Qy 239 EOKTNADALSN--SFT--KVIQONNEQSFVCTFTNANVPNSYFVAFSADVTVP----- 289  
 Db 239 DQKNNAQOFANVQSFITLDTLLENVEDA-----KKMQQAPANSYFVGYSVDVTGSGQET 292  
 Qy 290 ---NYKARFTWNGDEPSSRIANTNS-----ITDVSMTYSLAGTNTKYQFSNSY 338  
 Db 293 TIPWNNFAQRIWSSGNOPTVYATTTGGEOSTAKPLSDVSWIYSLAGTAKTLEFTFY 352  
 Qy 339 GPSTGYLYFPYKLVKADANNVGLQYKLN-NGNVQOVEFATSTSNANTTANPTPAVDEIK 397  
 Db 353 GPSTGYLYFPYKLVKAD--DVGQYKLNSEBELTPIIFBGGT---TTNPAATVETIN 406  
 Qy 398 VAKIVLSGLRFGQNTIELSVPTSGGNMKVAPMIGNIYLSNENNAK 445  
 Db 407 VAKVRLTGLAFGKNITIEFVSVP-----MSKVAPEMIGNMYITSSDYETNK 449

Mon Aug 25 09:49:57 2003

us-09-901-572a-3.rpt

Page 8

Search completed: August 14, 2003, 10:24:17  
Job time : 105 secs

---

THIS PAGE BLANK (USPTO)



C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D90011

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-2481 <KUR>

A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BA843253.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: fmcB(mrp)

Query Match 7.1%; Score 164.5; DB 2; Length 2481;  
Best Local Similarity 21.9%; Pred. No. 0.67;  
Matches 104; Conservative 66; Mismatches 173; Indels 131; Gaps 22;

```
QY 25 TONVTSREVSSVOLSEESTFYLCPPEVSGTVIRLEFGMSITKK--DANPNNGQTOL 81
Db 1150 TQDITAEIATAA-----NADVDNAVTVQANSIIEAANSQNDVDQAKTTGETSI 1196
QY 82 E-----AARMELTDLINAKAMTLASIQDYAKIEASLSAYSEAEFTNNNLN---- 127
Db 1197 DQVPTVKKKATPARNETAIILNKKIQEIQAATPDATDEKQAADA--EANTENGRANQAIS 1254
QY 128 --ATLEQLKMAKTNLESAINQANTDKTFDNEHPILVEAYKALKTTL---EQATNLEGL 182
Db 1255 AATTNAQVDEAKANAEALIN-AVTPKVVKKQAQKDEIDQLQATQTIVNNDQNAINEKE 1313
QY 183 S-----STAYNOIRNNL-----VDLYNKASSLIITKTLDPUNGTLILDSN-----E 222
Db 1314 AAIQOLATAVTDANKNITPAATDNGVDPAKDAKNSIQSTOP--ATAVKSNAKNEVDQA 1370
QY 223 ITTVNRNINNTLSTINEOKTNADALNSFIKKVIQNNESFVGFTTNANVOPSNYSPYAF 282
Db 1371 VTTONQALDNTTGAATTEKNAKADL-----VLAKAKAYODIL--NAQTN----- 1414
QY 283 SADVPVNYKYARRTVWNGDEPSSRIILANTNSITDVSWIYSLA---GTNTKYQFSFENV 338
Db 1415 --DVTQI-----KQAAVADIGGITADTTIKQVAKDELATKANEEKAL--- 1454
QY 339 GPSTGYLYFPYKLVKADAN-----NVGLQYKLNNGNVQOVEPATISTANNTTA--NPT 390
Db 1455 -----IAQTADATTEKEKQANQOVDAQLTQGN--ONIEVAQSIDDVNTAKDNAI 1501
QY 391 PAVDEIKYAKIVLSGLRFGQNTIELSVPTGSGNMKVAPMIGNTYLSSENNDAD 444
Db 1502 QALDPIDQASTDVKTAR-----AEILLTEM---QNKITEILLNNETTNEKGN 1546
```

Search completed: August 14, 2003, 10:22:28  
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 01:32:14 ; Search time 5026 Seconds  
(without alignments)  
10630.317 Million cell updates/sec

Title: US-09-901-572A-1  
Perfect score: 1306  
Sequence: 1 aaaaacacagatcgttcaat.....caaatcgtttacacagct 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBank:  
1: gb ba:\*  
2: gb hcg:\*  
3: gb in:\*  
4: gb om:\*  
5: gb ov:\*  
6: gb pat:\*  
7: gb ph:\*  
8: gb pl:\*  
9: gb pr:\*  
10: gb ro:\*  
11: gb scb:\*  
12: gb sy:\*  
13: gb un:\*  
14: gb vl:\*  
15: em ba:\*  
16: em fun:\*  
17: em hum:\*  
18: em in:\*  
19: em mu:\*  
20: em om:\*  
21: em or:\*  
22: em ov:\*  
23: em pat:\*  
24: em ph:\*  
25: em pl:\*  
26: em ro:\*  
27: em scb:\*  
28: em sy:\*  
29: em un:\*  
30: em vl:\*  
31: em hcg\_hum:\*  
32: em hcg\_inv:\*  
33: em hcg\_other:\*  
34: em hcg\_mus:\*  
35: em hcg\_pln:\*  
36: em hcg\_rtd:\*  
37: em hcg\_mam:\*  
38: em hcg\_vrt:\*  
39: em hcg\_hum:\*  
40: em hcg\_mus:\*  
41: em hcg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1280.4	98.0	2144	6 AR035279	AR035279 Sequence
2	1270	97.2	1387	6 AR035276	AR035276 Sequence
3	1264	96.8	1305	6 E09301	E09301 DNA encodin
4	1262.4	96.7	1387	6 E09301	E09301 DNA encodin
5	1179.4	90.3	302070	1 AE016968	AE016968 Mycoplasma
6	1086.8	83.2	2014	6 AR035278	AR035278 Sequence
7	1028	78.7	1152	6 AX65164	AX65164 Sequence
8	989.4	75.8	1015	6 E09896	E09896 Mycoplasma
9	986.2	75.5	1015	6 E09896	E09896 Mycoplasma
10	921.6	70.6	1082	6 AX65187	AX65187 Sequence
11	824.2	63.1	853	6 E02348	E02348 DNA sequenc
12	679.8	52.1	708	6 E02342	E02342 DNA sequenc
13	250.6	19.2	301042	1 AE016967	AE016967 Mycoplasma
14	249	19.1	4568	1 AE210770	AE210770 Mycoplasma
15	222.8	17.1	301903	1 AE016969	AE016969 Mycoplasma
16	217	16.6	7434	1 AE275312	AE275312 Mycoplasma
17	216.6	16.6	2369	6 AE035275	AE035275 Sequence
18	216.6	16.6	2369	6 E09895	E09895 Mycoplasma
19	215	16.5	3331	1 AF141940	AF141940 Mycoplasma
20	206.2	15.8	9720	1 MYCPMGAB	L28424 Mycoplasma
21	203.8	15.6	302070	1 AE016968	AE016968 Mycoplasma
22	199.8	15.3	3894	1 AY065985	AY065985 Mycoplasma
23	193.4	14.8	4715	1 AF032890	AF032890 Mycoplasma
24	191.8	14.7	8904	1 MGU90714	MGU90714 Mycoplasma
25	190.2	14.6	3897	1 AF053978	AF053978 Mycoplasma
26	173.2	13.3	2803	1 MYCAHOM	M83178 Mycoplasma
27	173.2	13.3	2809	1 S55216	S55216 pMGAL.2-hem
28	136.4	10.4	702	6 E02344	E02344 DNA sequenc
29	92	7.0	170627	2 AC125567	AC125567 Ratrus no
30	89	6.8	145265	2 AL935272	AL935272 Danio rer
31	85	6.5	110000	3 AC116305_1	Continuation (2 of
32	84.6	6.5	203138	2 BX530070	BX530070 Danio rer
33	82.4	6.3	94434	5 AL929250	AL929250 Zebrafish
34	82.2	6.3	155204	2 AC007926	AC007926 Trypanoso
35	80	6.1	154071	3 AC115598	AC115598 Dictyoste
36	79.2	6.1	153751	3 AC116551	AC116551 Dictyoste
37	79.2	6.1	225581	2 BX537105	BX537105 Danio rer
38	79	6.0	258920	2 BX510653	BX510653 Danio rer
39	78.8	6.0	178702	2 BX321884	BX321884 Danio rer
40	78.6	6.0	213544	2 BX510939	BX510939 Danio rer
41	78.2	6.0	594	6 E02345	E02345 DNA sequenc
42	77.4	5.9	335050	3 PFA929356	AL929356 Plasmodiu
43	77.2	5.9	67970	3 PFMAL1P3	AL031746 Plasmodiu
44	76.4	5.8	2167	3 PRSTRAPA	Z30339 P. reichenow
45	75.8	5.8	2426	8 SDU49822	U49822 Saccharomyc

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AR035279	AR035279	Sequence 9 from patent US 5871742.	AR035279	AR035279.1	GI:5951947	Unknown.	Unknown.	1 (bases 1 to 2144)	Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine
AR035279	AR035279	2144 bp	DNA	linear	PAT 29-SEP-1999	Unknown.	Unknown.	1 (bases 1 to 2144)	Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine

QY	1	AAAAACATCGATTGTTAATCTGAATCTCTTGCTTAAAAAAAACAATAATCTCTAACAA	60
Db	1	AAAAACATCGATTGTTAATCTGAATCTCTTGCTTAAAAAAAACAATAATCTCTAACAA	60
QY	61	AATCTTAATTAATAAGCCGTAATTAATAAATAATTAATAAATAGTTTTCTATC	120
Db	61	AATCTTAATTAATAAGCCGTAATTAATAAATAATTAATAAATAGTTTTCTATC	120
QY	121	AACCAAAATCTCTAGTAATAAAGCTTATTAATTTTATTTTATGATCCTTTTAGAT	180
Db	121	AACCAAAATCTCTAGTAATAAAGCTTATTAATTTTATTTTATGATCCTTTTAGAT	180
QY	181	ATAAATAATATCTTAATTAATTTCTATGAATAAGAAAAGATCATCTTAAGACTTATGTTG	240
Db	181	ATAAATAATATCTTAATTAATTTCTATGAATAAGAAAAGATCATCTTAAGACTTATGTTG	240
QY	241	TTTAGGTACAACATCTCTTTCTTAGACATGGGATTTCTAGCTGATGCTTAATCTAATAA	300
Db	241	TTTAGGTACAACATCTCTTTCTTAGACATGGGATTTCTAGCTGATGCTTAATCTAATAA	300
QY	301	GATGCAAAACCCAAATTAATAGCCAAATCCAAATTAAGAAAGCGCGAATGAGATTAAACAT	360
Db	301	GATGCAAAACCCAAATTAATAGCCAAATCCAAATTAAGAAAGCGCGAATGAGATTAAACAT	360
QY	361	CTAATCAATGCTAAACCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT	420
Db	361	CTAATCAATGCTAAACCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT	420
QY	421	AGTTTATCATCTGCTTATATGTAAGGTGAACAAGTTAACATTAATCAATCAATTAATGCAACATTA	480
Db	421	AGTTTATCATCTGCTTATATGTAAGGTGAACAAGTTAACATTAATCAATTAATGCAACATTA	480
QY	481	GAACAACATAAAATGCGCTAAATACTAATTTAGAAATCAGCATCAACCAAGCTAATACGAT	540
Db	481	GAACAACATAAAATGCGCTAATACTAATTTAGAAATCAGCATCAACCAAGCTAATACGAT	540
QY	541	AAAAAGCATTTTGATATGAAACACCCAAATTTAGTTGAAGCATACAAAGCATTAACAAAC	600
Db	541	AAAAAGCATTTTGATATGAAACACCCAAATTTAGTTGAAGCATACAAAGCATTAACAAAC	600
QY	601	ACTTTGAACAAGCTCTACTAACCCTGAAGGTTTGTCATCAACGCTTAATTAATCAATTT	660
Db	601	ACTTTGAACAAGCTCTACTAACCCTGAAGGTTTGTCATCAACGCTTAATTAATCAATTT	660
QY	661	CGCAATTAATTTAGTGATCTATACATAAAGCTAGTATTTAATTAATCTAATAACATAGAT	720
Db	661	CGCAATTAATTTAGTGATCTATACATAAAGCTAGTATTTAATTAATCTAATAACATAGAT	720
QY	721	CGAATTAATTTAGTGATCTATACATAAAGCTAGTATTTAATTAATCTAATAACATAGAT	780
Db	721	CGAATTAATTTAGTGATCTATACATAAAGCTAGTATTTAATTAATCTAATAACATAGAT	780
QY	781	AAATTAATACGTTATCAACTTAATTAAGAAACAAAGACTAATGCTAGATCTTAATTAATG	840
Db	781	AAATTAATACGTTATCAACTTAATTAAGAAACAAAGACTAATGCTAGATCTTAATTAATG	840
QY	841	TTTATTTAAAAAAGTGATCAAAATATATGAACAAAGTTTGTAGGCACTTTTACAAAGCT	900
Db	841	TTTATTTAAAAAAGTGATCAAAATATATGAACAAAGTTTGTAGGCACTTTTACAAAGCT	900
QY	901	AATGTTCAACCTTCAACATCAAGTTTTGTGCTTTAGTGTGATGTAACCCCGTCAT	960
Db	901	AATGTTCAACCTTCAACATCAAGTTTTGTGCTTTAGTGTGATGTAACCCCGTCAT	960

Db 241 TTAGTACAACATCCTTCTTAGCATGGGATTTCTAGCTGTATGCTATTAATAAAAA 30

```

Qy 301 GATGCAAAACCCAAATTAATGGCCAAACCCAAATTAGAGACGCCGGAATGAGTTAAACAGAT 360
Db 301 GAGCGAAACCCAAATTAATGGCCAAACCCAAATTAGAGAGCGCGAATGAGTTAACTGAT 360
Qy 361 CTAAATCAATGCTAAAGCCATGATGATTAAGTTCATCAACAGATTAAGCCAAATTAAGCT 420
Db 361 CTAAATCAATGCTAAAGCCATGATGATTAAGTTCATCAACAGATTAAGCTAAAGTTAGACT 420
Qy 421 AGTTATCATCTGCTTAATAGTAAAGCTGAAACAGTTAAACATTAATTAATGCAACATTA 480
Db 421 AGTTATCATCTGCTTAATAGTAAAGCTGAAACAGTTAAACATTAATTAATGCAACATTA 480
Qy 481 GAACAACCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 540
Db 481 GAACAACCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 540
Qy 541 AAAAGGACTTTTGAATTAATGAACACCCAAATTTAGTTGAAGCTTAACAAAGCTAAACCC 600
Db 541 AAAAGGACTTTTGAATTAATGAACACCCAAATTTAGTTGAAGCTTAACAAAGCTAAACCC 600
Qy 601 ACTTTAGAACAAACGCTGCTACTAAACCTTGGAAGGTTTGTCAATCACTGCTTAATCAAAAT 660
Db 601 ACTTTAGAACAAACGCTGCTACTAAACCTTGGAAGGTTTGTCAATCACTGCTTAATCAAAAT 660
Qy 661 CGCAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 720
Db 661 CGCAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 720
Qy 721 CCATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 780
Db 721 CCATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 780
Qy 781 AATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 840
Db 781 AATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 840
Qy 841 TTTATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 900
Db 841 TTTATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 900
Qy 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGTGATGATGAACACCCGCTCAAT 960
Db 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGTGATGATGAACACCCGCTCAAT 960
Qy 961 TATAATTAATGCAAGAGGACGTTTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAAT 1020
Db 961 TATAATTAATGCAAGAGGACGTTTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAAT 1020
Qy 1021 AACAGCAATAGTATCAAGATGTTTCTTGATTTTATGTTTATGCTGAACAAACAGAG 1080
Db 1021 AACAGCAATAGTATCAAGATGTTTCTTGATTTTATGTTTATGCTGAACAAACAGAG 1080
Qy 1081 TACCAATTAATGTTTATGCACTATGCTCATCACTGCTTATTTATTTTCCCTTATAG 1140
Db 1081 TACCAATTAATGTTTATGCACTATGCTCATCACTGCTTATTTATTTTCCCTTATAG 1140
Qy 1141 TTGGTTAAAGCACTGATGCTTAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAAT 1200
Db 1141 TTGGTTAAAGCACTGATGCTTAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAAT 1200
Qy 1201 GTTCAACAAGTTGATGCTTCCCACTTCAACTAGTGAACAAATTAATTAATCACTAATCAACT 1260
Db 1201 GTTCAACAAGTTGATGCTTCCCACTTCAACTAGTGAACAAATTAATTAATCACTAATCAACT 1260
Qy 1261 CCAGCAATGTTGATGATTAATGCTTAAATGCTTTTATCAAGTT 1306
Db 1261 CCAGCAATGTTGATGATTAATGCTTAAATGCTTTTATCAAGTT 1306

```

RESULT 3  
 LOCUS 117387 1305 bp DNA linear PAT 06-OCT-1996  
 DEFINITION Sequence 1 from patent US 5489430.

```

ACCESSION 117387
VERSION 117387.1 GI:1597742
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1305)
AUTHORS Saito,S., Ohkawa,S., Fujisawa,A., Iritani,Y. and Aoyama,S.
TITLE Polity mycoplasma antigen, gene thereof and recombinant vectors
JOURNAL Patent: US 5489430-A 1 06-FEB-1996;
FEATURES
BASE COUNT 499 a 208 c 184 g 408 t 6 others
ORIGIN
Query Match 96.8%; Score 1264; DB 6; Length 1305;
Best Local Similarity 98.3%; Pred. No. 2.2e-174;
Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
Qy 1 AAAAATCATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 1 AAAAATCATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy 61 AATCCTAAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 120
Db 61 AATCCTAAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 120
Qy 121 AACCAAAATCTCTAGTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 180
Db 121 AACCAAAATCTCTAGTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 180
Qy 181 AATAATATCTTATATTTCTATGATTAAGAAAGAAATCATCTTAAAGCATTTAGTTG 240
Db 181 AATAATATCTTATATTTCTATGATTAAGAAAGAAATCATCTTAAAGCATTTAGTTG 240
Qy 241 TTAGTCAACATCTTCTTCTAGCATTTGGGATTTCTGATGATGATGATGATGATGATGATGATGAT 300
Db 241 TTAGTCAACATCTTCTTCTAGCATTTGGGATTTCTGATGATGATGATGATGATGATGATGATGAT 300
Qy 301 GATGCAAAACCCAAATTAATGCGCAACCCAAATTAAGACAGCGGCAATGAGTTAAACAGAT 360
Db 301 GATGCAAAACCCAAATTAATGCGCAACCCAAATTAAGACAGCGGCAATGAGTTAAACAGAT 360
Qy 361 CTATTAATGCTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 CTATTAATGCTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 421 AGTTATCATCTGCTTAATAGTAAAGCTGAAACAGTTAAACATTAATTAATGCAACATTA 480
Db 421 AGTTATCATCTGCTTAATAGTAAAGCTGAAACAGTTAAACATTAATTAATGCAACATTA 480
Qy 481 GAACAACCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 540
Db 481 GAACAACCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 540
Qy 541 AAAAGGACTTTTGAATTAATGAACACCCAAATTTAGTTGAAGCTTAACAAAGCTAAACCC 600
Db 541 AAAAGGACTTTTGAATTAATGAACACCCAAATTTAGTTGAAGCTTAACAAAGCTAAACCC 600
Qy 601 ACTTTAGAACAAACGCTGCTACTAAACCTTGGAAGGTTTGTCAATCACTGCTTAATCAAAAT 660
Db 601 ACTTTAGAACAAACGCTGCTACTAAACCTTGGAAGGTTTGTCAATCACTGCTTAATCAAAAT 660
Qy 661 CGCAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 720
Db 661 CGCAATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 720
Qy 721 CCATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 780
Db 721 CCATTAATTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAAT 780

```

QY 781 AATAATACCTTATCACTATTAATGAAACAAAGACTATGATCATTTATCTATAGT 840  
Db 781 AATAATACCTTATCACTATTAATGAAACAAAGACTATGATCATTTATCTATAGT 840  
QY 841 TTTATTTAAAAAGTATGTAATAATGAAACAAAGTTTGTAGGACCTTTTCAACCGT 900  
Db 841 TTTATTTAAAAAGTATGTAATAATGAAACAAAGTTTGTAGGACCTTTTCAACCGT 900  
QY 901 AATGTTCAACCTTCAACACTGATTTGTTGCTTTTATGAGTCACTGATCAACCCGTCAT 960  
Db 901 AATGTTCAACCTTCAACACTGATTTGTTGCTTTTATGAGTCACTGATCAACCCGTCAT 960  
QY 961 TATTAATATGCAAGAAGACCGTTTGGATGATGAACTTCAAGTGAATTTCTTGA 1020  
Db 961 TATTAATATGCAAGAAGACCGTTTNNNATGATGAACTTCAAGTGAATTTCTTGA 1020  
QY 1021 AACAGAAATGATCAAGATGTTTCTTGATTTATGTTTATGTTGAGCAACAAACGAG 1080  
Db 1021 AACAGAAATGATCAAGATGTTTCTNNNATGTTATGTTTATGTTGAGCAACAAACGAG 1080  
QY 1081 TACCAATTTAGTTTATGCAACTATGATGATGATCACTGATTTATTTATTTCCCTATAG 1140  
Db 1081 TACCAATTTAGTTTATGCAACTATGATGATGATCACTGATTTATTTATTTCCCTATAG 1140  
QY 1141 TTGTTTAAAGCAGCTGATCTATTAACGTTGATTAACATACAAATTAATATGGAAT 1200  
Db 1141 TTGTTTAAAGCAGCTGATCTATTAACGTTGATTAACATACAAATTAATATGGAAT 1200  
QY 1201 GTTCAACAAGTTGATTTGCCACTTCACTAGTGAATTAATCTACAGCTATCTCACT 1260  
Db 1201 GTTCAACAAGTTGATTTGCCACTTCACTAGTGAATTAATCTACAGCTATCTCACT 1260  
QY 1261 CCAGCAGTTGATGATTAAGTTGCTAAATCGTTTATCAGTT 1306  
Db 1261 CCAGCAGTTGATGATTAAGTTGCTAAATCGTTTATCAGTT 1306

RESULT 4  
LOCUS E09301 1387 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding Mycoplasma TM-16.  
ACCESSION E09301  
VERSION E09301.1 GI:22025928  
KEYWORDS JP 1995133295-A/2.  
SOURCE Mycoplasma gallisepticum  
ORGANISM Mycoplasma gallisepticum  
Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
REFERENCE 1 (bases 1 to 1387)  
AUTHORS Mori,H., Saito,S., Okawa,S., Funato,H., Iritani,K., Aoyama,S. and Takahashi,K.  
TITLE NEW ANTIGEN PROTEIN, ITS GENE, RECOMBINANT BACULOVIRUS AND ITS USE  
JOURNAL Patent: JP 1995133295-A 2 23-MAY-1995.  
COMMENT NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1995133295-A/2  
PD 23-MAY-1995  
PF 27-AUG-1993 JP 1993213102  
PI MORI HAJIME, SAITO SHUJI, OKAWA SEISUKO, FUNATO HIRONO, PI  
IRITANI KOICHI,  
PI AOYAMA SHIGEMI, TAKAHASHI KIYOTO  
PC C07K14/30,A61K39/00,A61K39/00,C12N7/01,C12N15/31,C12P21/02, PC  
(C12P21/02,  
PC C12R1:92);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1..1387  
FT 202..1308 /organism='Mycoplasma gallisepticum' FT CDS  
FT Location/Qualifiers  
FT 1..1387

FEATURES  
source Location/Qualifiers  
1..1387

BASE COUNT 531 a 221 c 202 g 427 t 6 others  
ORIGIN  
Query Match 96.7%; Score 1262.4; DB 6; Length 1387;  
Best Local Similarity 98.2%; Pred. No. 3.7e-174;  
Matches 1283; Conservative 0; Mismatches 22; Indels 1; Gaps 1;  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
1 AAAAATCATGATGTTTATCTGATATCTTGTCTTAAAAAACAACAAATCTTCTACAA 60  
1 AAAAATCATGATGTTTATCTGATATCTTGTCTTAAAAAACAACAAATCTTCTACAA 60  
61 AATCTTAATTAATTAAGCCGTTAAATTAACCTAAAAAATTAATAAATGTTTCTTATC 120  
61 AATCTTAATTAATTAAGCCGTTAAATTAACCTAAAAAATTAATAAATGTTTCTTATC 120  
121 AACCAAAATCTCTAGTAAATTAAGCTTATTTATTTATTTATGATCTTTTAAAGAT 180  
121 AACCAAAATCTCTAGTAAATTAAGCTTATTTATTTATTTATTTATGATCTTTTAAAGAT 180  
181 AATAATATATCTTAATATCTTATGATTAAGAAAGAAATCACTTAAAGACTATTAAGTTG 240  
181 AATAATATATCTTAATATCTTATGATTAAGAAAGAAATCACTTAAAGACTATTAAGTTG 240  
241 TTAGTTCAACATCTTCTTATGATGGAATTTCTAGCTGATGCTATGCTATTAAGTAA 300  
241 TTAGTTCAACATCTTCTTATGATGGAATTTCTAGCTGATGCTATGCTATTAAGTAA 300  
301 GATGCAAAACCAAAATTAATGAGCAACCAATTAAGAGAGGCGCAATGAGTTAAACAGAT 360  
301 GATGCAAAACCAAAATTAATGAGCAACCAATTAAGAGAGGCGCAATGAGTTAAACAGAT 360  
361 CTAATCAATGCTTAAAGGATGATCACTTACCTTCACTACAGCTATGCCAAGATTGAAGCT 420  
361 CTAATCAATGCTTAAAGGATGATCACTTACCTTCACTACAGCTATGCCAAGATTGAAGCT 420  
421 AGTTATCATCTGTTTATGATGAGAGCTGAAACAGTTAACTAACTTAAAGCAACATTA 480  
421 AGTTATCATCTGTTTATGATGAGAGCTGAAACAGTTAACTAACTTAAAGCAACATTA 480  
481 GAAACAACTAAAAATGAGCTTAAATCTTAAATCTAGCACTCAACCAAGCTTAATACGAT 540  
481 GAAACAACTAAAAATGAGCTTAAATCTTAAATCTAGCACTCAACCAAGCTTAATACGAT 540  
541 AAAACGACTTTTATGATTAAGAACACCCAAATTTAGTTGATGATTAAGCACTTAAACC 600  
541 AAAACGACTTTTATGATTAAGAACACCCAAATTTAGTTGATGATTAAGCACTTAAACC 600  
601 AACTTGAACAAGAGTGAATCTTAACTTGAAGGTTGTCATCACTGCTTATTAATCAAT 660  
601 AACTTGAACAAGAGTGAATCTTAACTTGAAGGTTGTCATCACTGCTTATTAATCAAT 660  
661 CGCAATTAATTTAGTGAATCTTAACTTGAAGGTTGTCATCACTGCTTATTAATCAAT 720  
661 CGCAATTAATTTAGTGAATCTTAACTTGAAGGTTGTCATCACTGCTTATTAATCAAT 720  
721 CCATTAATGAGGGAAGGCTTTTATGATTTATGATTAAGTACTACAGTTAATCGGAATAT 780  
721 CCATTAATGAGGGAAGGCTTTTATGATTTATGATTAAGTACTACAGTTAATCGGAATAT 780  
781 AATAATATGTTATGATTAAGTAAAGCAAGCTATGATGATTAATCTTAAATAGT 840  
781 AATAATATGTTATGATTAAGTAAAGCAAGCTATGATGATTAATCTTAAATAGT 840  
841 TTTATTTAAAAAGTATGTAATAATGAAACAAAGTTTGTAGGACCTTTTCAACCGCT 900  
841 TTTATTTAAAAAGTATGTAATAATGAAACAAAGTTTGTAGGACCTTTTCAACCGCT 900  
901 AATGTTCAACCTTCAACACTGATTTGTTGCTTTTATGAGTCACTGATCAACCCGTCAT 960  
901 AATGTTCAACCTTCAACACTGATTTGTTGCTTTTATGAGTCACTGATCAACCCGTCAT 960

Oy 961 TATAATATGCAAGAGACCGTTTGATGTGTGATGAACCTTCAAGTAGAATCTTCA 1020  
 Db 961 TATAATATGCAAGAGACCGTTTNNATAGTGATGAACCTTCAAGTAGAATCTTCA 1020  
 Oy 1021 AACACGATAGTATACAGATGTTCTTGATTTATAGTTAGCTGAAACAACGAG 1080  
 Db 1021 AACACGATAGTATACAGATGTTCTTNNATTTATAGCTGAAACAACGAG 1080  
 Oy 1081 TACCAATTAGTTAGCAACTAGTGCATCACTGTTATTTATATTCCTTATAG 1140  
 Db 1081 TACCAATTAGTTAGCAACTAGTGCATCACTGTTATTTATATTCCTTATAG 1140  
 Oy 1141 TTGGTTAAGACGCTGATGCTAATACGTTGATTTACATCAATTAATATGGAAT 1200  
 Db 1141 TTGGTTAAGACGCTGATGCTAATACGTTGATTTACATCAATTAATATGGAAT 1200  
 Oy 1201 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCATTAATATACGCTAATCACT 1260  
 Db 1201 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCATTAATATACGCTAATCACT 1260  
 Oy 1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTT 1306  
 Db 1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTTATCAGGTT 1305

## RESULT 5

AB016968

LOCUS 302070 bp DNA linear BCT 09-JUN-2003  
 DEFINITION Mycoplasma gallisepticum strain R section 2 of 4 of the complete genome.

ACCESSION

AB016968

VERSION

AB016968.1

KEYWORDS

GI:31541264

SOURCE

ORGANISM

Mycoplasma gallisepticum R

Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

AUTHORS

Geary, S.J., Papazisi, L., Kutish, G., Gorton, T.S., Mahairas, G., Swartzell, S., Madan, A., Nguyen, D.K., Matham, P., Browning, G., Kamal, M., and Liao, X.

TITLE

The complete genome sequence of the avian pathogen Mycoplasma gallisepticum strain R

Microbiology (2003) In press

2 (bases 1 to 302070)

Geary, S.J., Papazisi, L., Kutish, G., Mahairas, G., Swartzell, S., Madan, A., Nguyen, D.K., Gorton, T.S., Matham, P., Browning, G., Mustafa, K., and Liao, X.

JOURNAL

REFERENCE

AUTHORS

TITLE

Direct Submision

Submitted (17-OCT-2002) Department of Pathobiology and Veterinary Sciences, and Center of Excellence for Vaccine Research, The University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT 06269-3089, USA

FEATURES

source

Location/Qualifiers

1. 302070

/organism="Mycoplasma gallisepticum R"

/mol\_type="genomic DNA"

/strain="R"

/db\_xref="taxon:23150"

79. 3588

/locus\_tag="MGA\_1011"

79. 3588

/locus\_tag="MGA\_1011"

/note="MGA\_1011; MGR\_224"

/codon\_start=1

/transl\_table=4

/product="conserved hypothetical"

/protein\_id="AAP5656.1"

/db\_xref="GI:31541265"

/translation="MMNNKKNNPKKPLVKKKGLDLDPESEIAVINDPHLKKPKKQPLLEKRAKKVQPVVYVGEKKVAKAKAVITDPRILKKSRLLITIGESLIGGITGPPLSFALINQIDKQDOSAPITRIDSPOFTHDNLSPFVNSYOLNKOSEINGIDISNORTLHDLPIASFMINKLVRLLNQQLFHNSYMLVSDIAYLIRKTY

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

DENKAFIPKLTDTLQITKSSDNNPLHVLITANEFNQTDSPTQFVFRFNORYKO  
 IOPKGINILKLSLSASNDIDPLKDPYISIDVDLAPGYINAVYKOLYANLYL  
 NMLNSNLEISWNSDSEYRSENFENNPNVSNILKTKIPGDPYQAOANVDLGEKML  
 RDLTPREISNIRKTFSENEPHSOEENNLIRDVKKMLIVNNNTKKXLYVSDAAAF  
 AKLPSYVFIKLNIRYANSKAILKULITDALSSEFTIYAKAVIEVLONDFAKULIL  
 TYEPFVAYSNLIDSLTKDVASADFQRLYEMLIRQLISNPEPKIKFEELIERL  
 GYRDDAGNLIPREFNTYLINTIPESRLLINDIYDOLINSVDFMAATANNNGSES  
 MSREFLRFNFELIONRSTLVNLFQSYNPEFDLITLFSNNFPLPOLKLFYNVAT  
 LDLEFKTALDILFLKNPEILSRDLKAKIIFWMLIKSLDQTNINKIKELIYKBNP  
 DNVLNOSPPTLTKKULIPDKLDINGLOIPIYGLSALFEPILPKIVDFPKYVDDK  
 RYERLKEKGYOQAINEGSSILNAGGIFSEKMACDPKAFSSVFDIMWKRLIKGVS  
 DNNETIOLANIEIPSTGNNDPNKVIATINGAGYFKVSGENVLDLPEELIPSNILN  
 QIVYQLVNDOEKDFSOIYONLKLHFRYSYRQPIFESDSKKIYQSIISQVQTLN  
 YSLNRKYLITFTRLILPGDGGKMLILQENNPYENVALFTYNAQSILNOYDRANED  
 KRRLKEOILMITDRBARVKGIDRDPKRETSILPQSLRMLDSILQIGVNSLYIHTN  
 MSPSTISERISLNTTEVKGLISKQVSNLAYETPLSASVQYLVLPYILDSNPN  
 PVFKKEISHLVSLFPGAGE"  
 3664. 5019  
 /locus\_tag="MGA\_1014"  
 3664. 5019  
 /locus\_tag="MGA\_1014"  
 /note="MGA\_1014; MGR\_225"  
 /codon\_start=1  
 /transl\_table=4  
 /product="unique hypothetical"  
 /protein\_id="AAP5657.1"  
 /db\_xref="GI:31541266"  
 /translation="MTPEYKKEIPDYDEGVDSIFHNDERTDPEDDQDHRGPRRN  
 RIKKNVYDQHDQTLTVSSKNNKFEPLKVRSTITVPIALWLLITLITVLAIVF  
 YLNKSPQOSESTYVSVALPLTIALAASVPIVTLGFMVNHKQIKYSLTERKSLF  
 YSIFPEKLYVESYFLGMIPLGLVILSVGSHLQNLDDSDVSDNNLMTQESISDLIDM  
 NGYENSRYLESKKVYLDOELORLANYEKKOKLAKNKLAVVSNALNPFKSKOLLYN  
 VNSNGYFASLIVBOVKSAITOLREBRKLRBRKRLKALISNANIDIEOPLIKETNO  
 LPNDGBESLFLNIDQPSRKNEGLENLNFDRIDIMISENLANDRAIYEQIKETNO  
 NLSRAPKVVKNQPTIKVKNVNNKKAALVQKQESDLTKQSNLITFNARLWNO  
 L"  
 5196. 5269  
 /product="tRNA-Trp"  
 /note="codon recognized: UGC; tRNA genes predicted using  
 PASTRNA software of the Institut Pasteur, France;  
 functional prediction of the tRNA is performed according  
 to Ushida C. and Muto"  
 6561. 7661  
 /gene="hacA"  
 /locus\_tag="MGA\_1017"  
 6561. 7661  
 /gene="hacB"  
 /locus\_tag="MGA\_1017"  
 /note="MGA\_1017; high affinity transport protein A (HacA)  
 [Mycoplasma gallisepticum]; MGR\_226"  
 /codon\_start=1  
 /transl\_table=4  
 /product="HacA"  
 /protein\_id="AAP5658.1"  
 /db\_xref="GI:31541267"  
 /translation="MSKTIKKLPLFSSPILSSAFVSSCAKANDANNQGTLOLVKDK  
 IHDANALINDYLKNSQYKOKLKNVTSNTSNTQGRKSIDPKSNLIDLAVSS  
 AIKTNDYITTKIOTLRSFRDTPPTFERAGTLDQKAKISLFPNFIPIRMTDNR  
 MWDNKKQPFYGGEDDLISFRKMLILGSDQTLAKIKEMAQKNMVFSSYGGGR  
 TSSNGRYTLPNLKKHNDPNLITISDQVGRGEICTNPVHIVPDDMSFATQNK  
 NKVNPYTFNPNDNTKVEILSLDPLALNIGVFNKRLTMDVINAVESIYELAKNNO  
 DTVGPTVGVNGYRIKDVNKEVQFIKANKIN"  
 7661. 8407  
 /gene="hacB"  
 /locus\_tag="MGA\_1018"  
 7661. 8407  
 /gene="hacB"  
 /locus\_tag="MGA\_1018"  
 /note="MGA\_1018; high affinity transport protein B (HacB)  
 [Mycoplasma gallisepticum]; MGR\_227"  
 /codon\_start=1  
 /transl\_table=4  
 /product="HacB"

gene  
/protein\_id="AAP56569.1"  
/db\_xref="GI:31541268"  
/translation="MSNVTNH KELDVISIKRKYKGNVDHINELKGSITLMLIKSG  
SGKITLINSIDKTLVSQGEIVPNDQLPFLSHKRIKLAKTIGLDQJLENLIDYI  
YENKIRISRKWYKFPFNITPISVLNKEVEVLTTLDLSHKLINIKDLSGGKQKQYD  
AKIFQPKKILIDEPITGLDINSEIIVEIKLSEQFAPVITISIHLEIVKKEAD  
KILIIQNKPKPILINNSADLNIIDELIRKKYET"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

CDS  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

gene  
/protein\_id="AAP56570.1"  
/db\_xref="GI:31541269"  
/translation="WRLEKNMNRHKPFROLKINKTVEFTVFLICITSAIVKGIN  
AAGLVAKNINISNFPNOSAVNKSLSLTKSPQMLTIKTAATGTITGILIGLIGS  
SLRINNNPALTAVKGLIFPSPFPVAVIPLNPNFSNDSALITISWSPMLNNAKTI  
ADYINSEVQEPKFKVNSQNHLSPTINNIFIDIKSKILVLFVAVSLESNRCVTTIISA  
VGLIIGOLINDPISINDDFASTLIPLIVAVPELLINMGALVLEFQYLVARKSPYKT  
NINKLTKIKVYLVIFILITPISYISFISRLRVNLITPDPQRLFNPKSPFAT  
TTPNNPRLNITLITLOALILIGIVPESLIPALINRVSLEPKALPLIKTI  
PLIIVFRILPNLFPNSGITIVPISGVPSLAKKIYVLININMTVYVLSKSKVTN  
FELIIVYIIPISRKDLVTVQLEFESLIRTTITLGAVGTSVIGQLDIYIYKRSIENL  
GSYVISIMVFEQVIDLSIMVRKKFPFTNK"  
10143. .11567  
/gene="pros"  
/locus\_tag="MGA\_1022"  
10143. .11567  
/gene="pros"  
/locus\_tag="MGA\_1022"  
/note="MGA\_1022: [U] COG0442 Prol1-trNA synthetase;  
MGR\_229"  
/codon\_start=1  
/transl\_table=4  
/product="pros"  
/protein\_id="AAP56571.1"  
/db\_xref="GI:31541270"  
/translation="MAFKKNLASKSDONI SDWYTEVYSKGNLIHYSNIKPMSEFLNG  
KIMLOAIKELTKEPDRIILELCLPSLISONDFLEKGLGAPBLFTVTSNET  
EXKILAPTESEIAFNLMBETLANYKQLPIKHNQTSYVRKKNTPILRNSBEFFMLP  
HSGFEFBNDSFAVDIMKLYORFIKQIDICPLVAGEKTEKFAKGTATVTEIIMP  
DQALOSASRSHLSIIONFSKAPRIPRYQTNQNNYONVFSAGVSTRIGAILMHGDD  
GLVFPKPAVPHIISLANCIPTDDINOELNAKTEILANKYSOKRVHLANNNHSTEBIIN  
SOLRGSCVALLGPNDLKNEIVFTIDRIKQKQPIINLDHLDQKLELPSFPDQKQK  
AKAVPTKVDVQATFEPEEQKIASGKPFVRVYCNEDYERKQIKETGTAASSCIITKID  
EQOERCFISNKKAKVETIFARSY"  
11768. .12190  
/locus\_tag="MGA\_1027"  
11768. .12190  
/locus\_tag="MGA\_1027"  
/note="MGA\_1027; MGR\_230"  
/codon\_start=1  
/transl\_table=4  
/product="unique hypothetical"  
/protein\_id="AAP56572.1"  
/db\_xref="GI:31541271"  
/translation="MGMPFYKIAKKIPIHTPVYKQHILINROBERSISNMRIINDAKKY  
LEHYDYRKATIERKQIISVFNHIDGVDNQSOMGYSVDLIDATIKRLKKEKDLIDIV  
FLRKIKRNEFNYSQAFYRLNKKACCSLYSLGVITPN"  
12194. .13276  
/locus\_tag="MGA\_1029"  
12194. .13276  
/locus\_tag="MGA\_1029"  
/note="MGA\_1029; MGR\_231"  
/codon\_start=1

CDS  
/locus\_tag="MGA\_1029"  
12194. .13276  
/locus\_tag="MGA\_1029"  
/note="MGA\_1029; MGR\_231"  
/codon\_start=1

gene  
/protein\_id="AAP56573.1"  
/db\_xref="GI:31541272"  
/translation="MSNVTNH KELDVISIKRKYKGNVDHINELKGSITLMLIKSG  
SGKITLINSIDKTLVSQGEIVPNDQLPFLSHKRIKLAKTIGLDQJLENLIDYI  
YENKIRISRKWYKFPFNITPISVLNKEVEVLTTLDLSHKLINIKDLSGGKQKQYD  
AKIFQPKKILIDEPITGLDINSEIIVEIKLSEQFAPVITISIHLEIVKKEAD  
KILIIQNKPKPILINNSADLNIIDELIRKKYET"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

CDS  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

gene  
/protein\_id="AAP56574.1"  
/db\_xref="GI:31541273"  
/translation="WRLEKNMNRHKPFROLKINKTVEFTVFLICITSAIVKGIN  
AAGLVAKNINISNFPNOSAVNKSLSLTKSPQMLTIKTAATGTITGILIGLIGS  
SLRINNNPALTAVKGLIFPSPFPVAVIPLNPNFSNDSALITISWSPMLNNAKTI  
ADYINSEVQEPKFKVNSQNHLSPTINNIFIDIKSKILVLFVAVSLESNRCVTTIISA  
VGLIIGOLINDPISINDDFASTLIPLIVAVPELLINMGALVLEFQYLVARKSPYKT  
NINKLTKIKVYLVIFILITPISYISFISRLRVNLITPDPQRLFNPKSPFAT  
TTPNNPRLNITLITLOALILIGIVPESLIPALINRVSLEPKALPLIKTI  
PLIIVFRILPNLFPNSGITIVPISGVPSLAKKIYVLININMTVYVLSKSKVTN  
FELIIVYIIPISRKDLVTVQLEFESLIRTTITLGAVGTSVIGQLDIYIYKRSIENL  
GSYVISIMVFEQVIDLSIMVRKKFPFTNK"  
10143. .11567  
/gene="pros"  
/locus\_tag="MGA\_1022"  
10143. .11567  
/gene="pros"  
/locus\_tag="MGA\_1022"  
/note="MGA\_1022: [U] COG0442 Prol1-trNA synthetase;  
MGR\_229"  
/codon\_start=1  
/transl\_table=4  
/product="pros"  
/protein\_id="AAP56575.1"  
/db\_xref="GI:31541274"  
/translation="MAFKKNLASKSDONI SDWYTEVYSKGNLIHYSNIKPMSEFLNG  
KIMLOAIKELTKEPDRIILELCLPSLISONDFLEKGLGAPBLFTVTSNET  
EXKILAPTESEIAFNLMBETLANYKQLPIKHNQTSYVRKKNTPILRNSBEFFMLP  
HSGFEFBNDSFAVDIMKLYORFIKQIDICPLVAGEKTEKFAKGTATVTEIIMP  
DQALOSASRSHLSIIONFSKAPRIPRYQTNQNNYONVFSAGVSTRIGAILMHGDD  
GLVFPKPAVPHIISLANCIPTDDINOELNAKTEILANKYSOKRVHLANNNHSTEBIIN  
SOLRGSCVALLGPNDLKNEIVFTIDRIKQKQPIINLDHLDQKLELPSFPDQKQK  
AKAVPTKVDVQATFEPEEQKIASGKPFVRVYCNEDYERKQIKETGTAASSCIITKID  
EQOERCFISNKKAKVETIFARSY"  
11768. .12190  
/locus\_tag="MGA\_1027"  
11768. .12190  
/locus\_tag="MGA\_1027"  
/note="MGA\_1027; MGR\_230"  
/codon\_start=1  
/transl\_table=4  
/product="unique hypothetical"  
/protein\_id="AAP56576.1"  
/db\_xref="GI:31541275"  
/translation="MGMPFYKIAKKIPIHTPVYKQHILINROBERSISNMRIINDAKKY  
LEHYDYRKATIERKQIISVFNHIDGVDNQSOMGYSVDLIDATIKRLKKEKDLIDIV  
FLRKIKRNEFNYSQAFYRLNKKACCSLYSLGVITPN"  
12194. .13276  
/locus\_tag="MGA\_1029"  
12194. .13276  
/locus\_tag="MGA\_1029"  
/note="MGA\_1029; MGR\_231"  
/codon\_start=1

CDS  
/locus\_tag="MGA\_1029"  
12194. .13276  
/locus\_tag="MGA\_1029"  
/note="MGA\_1029; MGR\_231"  
/codon\_start=1

gene  
/protein\_id="AAP56577.1"  
/db\_xref="GI:31541276"  
/translation="MSNVTNH KELDVISIKRKYKGNVDHINELKGSITLMLIKSG  
SGKITLINSIDKTLVSQGEIVPNDQLPFLSHKRIKLAKTIGLDQJLENLIDYI  
YENKIRISRKWYKFPFNITPISVLNKEVEVLTTLDLSHKLINIKDLSGGKQKQYD  
AKIFQPKKILIDEPITGLDINSEIIVEIKLSEQFAPVITISIHLEIVKKEAD  
KILIIQNKPKPILINNSADLNIIDELIRKKYET"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

CDS  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
8379. .9998  
/gene="hacC"  
/locus\_tag="MGA\_1019"  
/note="MGA\_1019: high affinity transport protein C (HacC)  
[Mycoplasma gallisepticum]: an alternative start codon may  
be 27 nucleotides upstream: "tta aca att cgg cgg atc tta  
ata tcg atg / 'LTLIKLISM' ; MGR\_228"

gene  
/protein\_id="AAP56578.1"  
/db\_xref="GI:31541277"  
/translation="WRLEKNMNRHKPFROLKINKTVEFTVFLICITSAIVKGIN  
AAGLVAKNINISNFPNOSAVNKSLSLTKSPQMLTIKTAATGTITGILIGLIGS  
SLRINNNPALTAVKGLIFPSPFPVAVIPLNPNFSNDSALITISWSPMLNNAKTI  
ADYINSEVQEPK

Query Match	90.3%	Score 1179.4	DB 1	Length 3020PTD
Best Local Similarity	95.3%	Pred. No. 7,2e-163		
Matches 1249	Conservative	0	Mismatches 56	Indels 6
				Gaps 3
QY	1	AAAAACATCGATGGTAACTCTGATATCTTGGCTTAAAAAACAACAAATCTTCAACAA	60	
Db	189913	AGAACATCGATGGTAACTCTGATATCTTGGCTTAAAAAACAACAAATCTTCAACAA	189971	
QY	61	AATCTTAAATTAATTAAGCCGTTAAATTAACATAAAATTAATAAAATGGTTTTCTATC	120	
Db	189972	AATCTTAAATTAATTAAGCCGTTAAATTAACATAAAATTAATAAAATGGTTTTCTATC	190031	
QY	121	AACCAAAATCTCTAGTAATTAAGCGTTATTAATTTATTTAGTCAATCTTTAAGA	179	
Db	190032	AACCAAAATCTTATAGTAATTAAGCGTTATTAATTTATTTAGTCAATCTTTAAGA	190091	
QY	180	TA---TAAATATATCTTAAATATTTCTATGATAAAGAAAGATCATCTTAAAGCTATTA	235	
Db	190092	TATATCTATATATATCTTAAATATTTCTATGATAAAGAAAGATCATCTTAAAGCTATTA	190151	
QY	236	GTTTGTAGGTACAACATCTTCTTGAAGTGGGATTTCTAGCTGATGCTATTTCTA	295	
Db	190152	GTTTGTAGGTACAACATCTTCTTGAAGTGGGATTTCTAGCTGATGCTATTTCTA	190211	
QY	296	AAAAAGATGAAACCCCAATATATGAGCCCAACCCCAATTAGAACAGCGGAATGAGTTAA	355	
Db	190212	AAAAAGAGCCAAATCCAAATATATGAGCCCAACCCCAATTAGAACAGCGGAATGAGTTAA	190271	
QY	356	CAGATCTATATCATGCTTAAAGCGATGACATTAGCTTCACTACAGACCTATGCCAAGATTG	415	
Db	190272	CTGATCTATATCAATGCTTAAAGCGATGACATTAGCTTCACTACAGAGATTATGCCAAGATTG	190331	
QY	416	AAGCTAGTTTATCATCTGCTTATATGTAAGTGAAGCTGAAACAGTTAACATTAACCTTATGCA	475	
Db	190332	AAGCTAGTTTATCATCTGCTTATATGTAAGTGAAGCTGAAACAGTTAACATTAACCTTATGCA	190391	
QY	476	CATTAGAACATTAATAATGGCTAAATACTAATTTGAATTCAGCCATCAACCAAGTATTA	535	
Db	190392	CATTAGAACATTAATAATGGCTAAATACTAATTTGAATTCAGCCATCAACCAAGTATTA	190451	
QY	536	CGGATTAACACATTTTGTATATGTAAGCAACCCAAATTTTGTGAAGCATACAAAGCACTTA	595	
Db	190452	CGGATTAACACATTTTGTATATGTAAGCAACCCAAATTTTGTGAAGCATACAAAGCACTTA	190511	
QY	596	AAACCACTTTAGAACACAGTGTCTATCACTTGAAGGTTTGTCACTCACTGCTTATATC	655	
Db	190512	AAACCACTTTATCAACAGTGTCTATCACTTGAAGGTTTGTCACTCACTGCTTATATC	190571	
QY	656	AAATTCGCAATTAATTTAGTGTATCTATACATTAAGCTAGTAAAGTAAATTAACATAACAC	715	
Db	190572	AGATTCGCAATTAATTTAGTGTATCTATACATTAAGCTAGTAAAGTAAATTAACATAACAC	190631	
QY	716	TAGATCCACTAAATGGGGGAAAGCTTTTGAATTTCTATGAGTTTCTACAGTTAAATCGGA	775	
Db	190632	TAGATCCACTAAATGGGGGAAAGCTTTTGAATTTCTATGAGTTTCTACAGCTAAATTAAGA	190691	
QY	776	ATATTAATTAATTCGTTATCACTATTAATTAAGAACAAAGCTTAATGCTATGATTAATCTGA	835	
Db	190692	ATATTAATTAATTCGTTATCACTATTAATTAAGAACAAAGCTTAATGCTATGATTAATCTGA	190751	
QY	836	ATAGTTTATTAATAAAAGGATGTCAAAATTAATGAACAAAGTTTGTAGGACCTTTTACAA	895	
Db	190752	ATAGTTTATTAATAAAAGGATGTCAAAATTAATGAACAAAGTTTGTAGGACCTTTTACAA	190811	



QY 896 ACGTAATGTTCAACCTTCAACTACAGTTTGTGCTTTAGTGTGATGTAACACCCG 955  
DB 190812 ACACTAATGTTCAACCTTCAACTACAGTTTGTGCTTTAGTGTGATGTAACACCTG 190871  
QY 956 TCAATTAATTAATGTAAGAAAGACCGTTTGAATGCTGTAACCTTCAAGTAGAATTC 1015  
DB 190872 TTAATTAATTAATGTAAGAAAGACCGTTTGAATGCTGTAACCTTCAAGTAGAATTC 190931  
QY 1016 TTGCAACACGAAATGTAATCAAGATGTTCTGATTTAGTTAGTGTGAAACA 1075  
DB 190932 TTGCAACACGAAATGTAATCAAGATGTTCTGATTTAGTTAGTGTGAAACA 190991  
QY 1076 CGAAGTACCAATTTAGTTTGAACAATGCTGCTCACTCACTGTTATTTATTTCCCTT 1135  
DB 190992 CGAAGTACCAATTTAGTTTGAACAATGCTGCTCACTCACTGTTATTTATTTCCCTT 191051  
QY 1136 ATTAAGTTGTTAAACAGCTGATGCTAATTAACGTTGATTAACAATTAATTAATG 1195  
DB 191052 ATTAAGTTGTTAAACAGCTGATGCTAATTAACGTTGATTAACAATTAATTAATG 191111  
QY 1196 GAAATGTTCAACAGTTGATTTGCACTTCAACTAGCAAAATTAATTAATTAATG 1255  
DB 191112 GAAATGTTCAACAGTTGATTTGCACTTCAACTAGCAAAATTAATTAATTAATG 191171  
QY 1256 CAATCCAGCAGTTGATGATTAATTAAGTTGCTAAATCGTTTATCAGGTT 1306  
DB 191172 CAATCCAGCAGTTGATGATTAATTAAGTTGCTAAATCGTTTATCAGGTT 191222

RESULT 6  
AR035278 AR035278 2014 bp DNA linear PAT 29-SEP-1999  
LOCUS Sequence 7 from patent US 5871742.  
DEFINITION AR035278  
ACCESSION AR035278.1 GI:5951946  
VERSION AR035278.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2014)  
AUTHORS Salton, S., Ohkawa, S., Saeki, S., Ohsawa, I., Funato, H., Iritani, Y., Aoyama, S. and Takahashi, K.  
TITLE Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine  
PATENT: US 5871742-A 7 16-FEB-1999;  
FEATURES  
source 1..2014  
location/Qualifiers  
BASE COUNT 757 a 324 c 318 g 615 t  
ORIGIN

Query Match 83.2%; Score 1086.8; DB 6; Length 2014;  
Best Local Similarity 96.7%; Pred. No. 1e-148;  
Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 150 TTTATTTTATTTTATG-CATCTTTTAAAGATATATATCTTAATATCTAGATA 208  
DB 1 TTTATTTTATTTTATGTAATCTTTTAAATATATATATTTTAAATATCTAGATA 60  
QY 209 AGAAAAGATCATCTTAAGAATACTATAGTTGTAAGTACAACATCTTTCTAGCATG 268  
DB 61 AAAAAAGATCATCTTAAGAATACTATAGTTGTAAGTACAACATCTTTCTAGCATG 120  
QY 269 GGATTTCTAGCTGATGCTATCTATTAATAAAGATGCAACCCAAATATGCGCAACCC 328  
DB 121 GGATTTCTAGCTGATGCTATCTATTAATAAAGATGCAACCCAAATATGCGCAACCC 180  
QY 329 AATTGAAGACGCGGAGATGAGTTTACAGATCTTAATCAATGCTTAAAGCGATGACATTG 388  
DB 181 AATTGAAGACGCGGAGATGAGTTTACAGATCTTAATCAATGCTTAAAGCGATGACATTG 240  
QY 389 CTTCATCAAGAACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG 448

DB 241 CTTCATCAAGAACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG 300  
QY 449 AAAAGTTAAACAATTAACCTTAAATGCAACATTAGAAACAATAATGGCTAAACTAAT 508  
DB 301 AAAAGTTAAACAATTAACCTTAAATGCAACATTAGAAACAATAATGGCTAAACTAAT 360  
QY 509 TAGAATGAGCATCAACCAAGCTAATGCGATAAAGACCTTTGATTAATGACACCCA 568  
DB 361 TAGAATGAGCATCAACCAAGCTAATGCGATAAAGACCTTTGATTAATGACACCCA 420  
QY 569 ATTTAGTTGAAGCATTAACAAGCACTAAACCACTTTAGAACCAAGTGTACTAACCTTG 628  
DB 421 ATTTAGTTGAAGCATTAACAAGCACTAAACCACTTTAGAACCAAGTGTGTACTAACCTTG 480  
QY 629 AAGTTTGTCACTCAACGCTTAAATGCAATTCGCAATTAATTTGATGATCTATACATA 688  
DB 481 AAGTTTGTCACTCAACGCTTAAATGCAATTCGCAATTAATTTGATGATCTATACATA 540  
QY 689 AAGCTAGATTTAATTAATACTAAACACTAGATCCACTAAATGGGGAAACGCTTTAGATT 748  
DB 541 AAGCTAGATTTAATTAATACTAAACACTAGATCCACTAAATGGGGAAACGCTTTAGATT 600  
QY 749 CTAAATGATTTACTACAGTTAATCGAATATTAATTAATTAATGCTTATCACTAATGAAAC 808  
DB 601 CTAAATGATTTACTACAGTTAATCGAATATTAATTAATTAATGCTTATCACTAATGAAAC 660  
QY 809 AAAAGCTAATGCTGATGCAATATCTAATAGTTTATTAATAAAGTATTAATAATG 868  
DB 661 AAAAGCTAATGCTGATGCAATATCTAATAGTTTATTAATAAAGTATTAATAATG 720  
QY 869 AACAAAGTTTGTAGGAGCTTTTACAAACGCTTAATGTTCAACCTTCAACCTACAGTTTGG 928  
DB 721 AACAAAGTTTGTAGGAGCTTTTACAAACGCTTAATGTTCAACCTTCAACCTACAGTTTGG 780  
QY 929 TTGCTTTTATGCTGATGTAACACCGCTCAATTAATTAATTAATGCAAGAAAGACCGTTTGA 988  
DB 781 TTGCTTTTATGCTGATGTAACACCGCTCAATTAATTAATTAATTAATGCAAGAAAGACCGTTTGA 840  
QY 989 ATGCTGATGAACCTTCAAGTGAATTTCTTGCAACAGCAATGATTAACAGATGTTTCTT 1048  
DB 841 ATGCTGATGAACCTTCAAGTGAATTTCTTGCAACAGCAATGATTAACAGATGTTTCTT 900  
QY 1049 GGATTTTATGTTTATGCTGGAACAACAAGCAAGTACCAATTTAGTTTGAACAATGATGTC 1108  
DB 901 GAATTTTATGTTTATGCTGGAACAACAAGCAAGTACCAATTTAGTTTGAACAATGATGTC 960  
QY 1109 CATCACTGCTTATTTATTTCCCTTAATAAGTTGTTAAAGCAGCTGATGCTAATAACG 1168  
DB 961 CATCACTGCTTATTTATTTCCCTTAATAAGTTGTTAAAGCAGCTGATGCTAATAACG 1020  
QY 1169 TTGATTTACAAATTAATAATTAATGAATGTTCAACAAGTGTGATGCTGCACTTCA 1228  
DB 1021 TTGATTTACAAATTAATAATTAATGAATGTTCAACAAGTGTGATGCTGCACTTCA 1080  
QY 1229 CTAGTGAATAATTAATTAATGCTAATCAACTCCAGAGGTTGATGATTAAGTTGCTA 1288  
DB 1081 CTAGTGAATAATTAATTAATGCTAATCAACTCCAGAGGTTGATGATTAAGTTGCTA 1140  
QY 1289 AAATCGTTTATCAAGTT 1306  
DB 1141 AAATCGTTTATCAAGTT 1158

RESULT 7  
AX665164 AX665164 1152 bp DNA linear PAT 26-MAR-2003  
LOCUS Sequence 1 from Patent EP1275716.  
DEFINITION AX665164  
ACCESSION AX665164  
VERSION AX665164.1 GI:29290294  
KEYWORDS  
SOURCE Mycoplasma gallisepticum  
ORGANISM Mycoplasma gallisepticum

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Zeon Corporation (JP)  
 Location/Qualifiers  
 1.1152  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:2096"  
 /note="TMM-1 gene (after EcorI)"

BASE COUNT  
 437 a 190 c 188 g 337 t

ORIGIN

Query Match  
 Best Local Similarity 78.7%; Score 1028; DB 6; Length 1152;  
 Matches 1028; Conservative 0; Pred. No. 4.2e-140; Mismatches 0; Indels 0; Gaps 0;

279 CTGTATGCTATTAACCAAGATGCAACCAATATGAGCCAAATGAGC 338  
 6 CTGTATGCTATTAACCAAGATGCAACCAATATGAGCCAAATGAGC 65  
 339 AGCGCGATGAGTAAACAGATCTAATGCTAAAGGATGACATTAGCTTCA 398  
 66 AGCGCGATGAGTAAACAGATCTAATGCTAAAGGATGACATTAGCTTCA 125  
 399 AGAGTATGCCAGATGAACTGTTATCATCTGTTATAGGAGTGAACAGTTAA 458  
 126 AGAGTATGCCAGATGAACTGTTATCATCTGTTATAGGAGTGAACAGTTAA 185  
 459 CAATTAACCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAA 518  
 186 CAATTAACCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAA 245  
 519 CATCAACCAAGCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAA 578  
 246 CATCAACCAAGCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAA 305  
 579 AGAGTATGCCAGATGAACTGTTATCATCTGTTATAGGAGTGAACAGTTAA 638  
 306 AGAGTATGCCAGATGAACTGTTATCATCTGTTATAGGAGTGAACAGTTAA 365  
 639 ATCAAGCTTATTAATCAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 698  
 366 ATCAAGCTTATTAATCAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 425  
 699 TTTAATTAACCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 758  
 426 TTTAATTAACCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 485  
 759 TACTACAGTATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 818  
 486 TACTACAGTATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 545  
 819 TCGTATGCTATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 878  
 546 TCGTATGCTATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 605  
 879 TGTAGGAGCTTTTAAACAGCTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 938  
 606 TGTAGGAGCTTTTAAACAGCTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 665  
 939 TGTAGGAGCTTTTAAACAGCTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 998  
 666 TGTAGGAGCTTTTAAACAGCTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 725  
 999 ACCCTCAAGTAGATTTCTGCAACAGCAATAGTATCAAGATGTTTCTTGATTTATAG 1058  
 726 ACCCTCAAGTAGATTTCTGCAACAGCAATAGTATCAAGATGTTTCTTGATTTATAG 785  
 1059 TTAGCTGAGCAACAGCAATAGTATCAAGATGTTTCTTGATTTATAG 1118

DB  
 786 TTAGCTGAGCAACAGCAATAGTATCAAGATGTTTCTTGATTTATAG 845  
 1119 TTATTAATTAATTCCTTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 1178  
 846 TTATTAATTAATTCCTTTAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAA 905  
 1179 ATCAATTAATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 1238  
 906 ATCAATTAATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 965  
 1239 TATATACAGCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 1298  
 966 TATATACAGCTTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 1025  
 1299 ATCAGGTT 1306  
 1026 ATCAGGTT 1033

## RESULT 8

E09896  
 LOCUS  
 DEFINITION  
 Mycoplasma gallisepticum TM-1 gene.  
 E09896  
 ACCESSION  
 E09896.1 GI:22026524  
 VERSION  
 JP 1995236498-A/2.  
 KEYWORDS  
 Mycoplasma gallisepticum  
 SOURCE  
 ORGANISM  
 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 DETECTION OF MYCOPLASMA GALLISEPTICUM AND NEW DNA TO BE USED  
 Patent: JP 1995236498-A 2 12-SEP-1995;  
 NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
 OS Mycoplasma gallisepticum  
 PN JP 1995236498-A/2  
 PD 12-SEP-1995  
 PF 25-FEB-1994 JP 1994052764  
 PI SAITO SHUJI, KYO TSUGUO, OKAWA SETSURO, IRIHARA KOICHI PC  
 C12Q1/68, C12N15/09, (C12Q1/68, C12R1:35), (C12N15/09, C12R1:35); CC  
 CC topology: linear;  
 FH key  
 FH Location/Qualifiers

## COMMENT

FT source 1.1015  
 FT /organism="Mycoplasma gallisepticum" FT CDS  
 FT 202..987  
 FT /product="TM-1 gene product"  
 FT misc\_feature 718..741  
 FT /note="this region is favorable for FT  
 FT hybridization probe to  
 FT detect TM-1 gene".  
 FT Location/Qualifiers  
 1.1015  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:2096"

## BASE COUNT

400 a 162 c 138 g 315 t

## ORIGIN

Query Match  
 Best Local Similarity 75.8%; Score 989.4; DB 6; Length 1015;  
 Matches 999; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

1 AAAAATCATGATGTTATCTGATATCTTTGCTTAAACAAATGCTTAAACAA 60  
 1 AAAAATCATGATGTTATCTGATATCTTTGCTTAAACAAATGCTTAAACAA 60  
 61 AATCCATTAATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 120  
 61 AATCCATTAATTAATGCAACATTAAGACACTAAATGCTAAATGCTAAATGCTAAATGCTAA 120

Qy	121	AACGAAATTCCTGCTATATAAAGCCTATTATATTTATTTTGTGCACTTTAAGAT	180
Dp	121	AACCAAAATTCCTGATATATAAGCCTATTATTTATTTATTTTGTGCACTTTAAGAT	180
Qy	181	ATTAATATATCTTAATATTTCTATGATATAAGAAAGATCATCTTAAGAATCTTAATGTTTG	240
Dp	181	ATTAATATATCTTAATATTTCTATGATATAAGAAAGATCATCTTAAGAATCTTAATGTTTG	240
Qy	241	TTAGGTACAAACATCCTTTCTTAGCATTTGGGATTTCTAGCTGATGCTATTTAATAAAAA	300
Dp	241	TTAGGTACAAACATCCTTTCTTAGCATTTGGGATTTCTAGCTGATGCTATTTAATAAAAA	300
Qy	301	GATGAAACCCAAATTAATGCGCAAAACCCAAATTAGAAAGCGCGGAATGAGTTAACAT	360
Dp	301	GACGAAACCCAAATTAATGCGCAAAACCCAAATTAGAAAGCGCGGAATGAGTTAACAT	360
Qy	361	CTAATCAATGCTTAAGAGCATGATCATTTAGCTTCACTACAAAGATTAAGCCAAAGTTAGAGCT	420
Dp	361	CTAATCAATGCTTAAGAGCATGATCATTTAGCTTCACTACAAAGATTAAGCTTAAGTTAGAGCT	420
Qy	421	AGTTTATCATCTGCTTAATAGTGAAGCTGAAACAGTTAACATTAACCTTAATGSCAACATTA	480
Dp	421	AGTTTATCATCTGCTTAATAGTGAAGCTGAAACAGTTAACATTAACCTTAATGSCAACATTA	480
Qy	481	GAAACAATAAATGCGCTAAAACTAATTTGAATCAGCCTACACCAAGCTTAATACGAT	540
Dp	481	GAAACAATAAATGCGCTAAAACTAATTTGAATCAGCCTACACCAAGCTTAATACGAT	540
Qy	541	AAAACGACTTTTGATTAATGAACACCCAAATTAGTTGAAGCATACAAAGCATTAATAAAC	600
Dp	541	AAAACGACTTTTGATTAATGAACATCCAAATTTAGTTGAAGCATACAAAGCATTAATAAAC	600
Qy	601	ACTTTAGAACAAACGTCCTACTAACCTTGAGGTTTGTCACTCAACGCTTAATATCAAT	660
Dp	601	ACTTTAGAACAAACGTCCTACTAACCTTGAGGTTTGTCACTCAACGCTTAATATCAAGAT	660
Qy	661	CGCAATATTTAGTGGATCTATACAAATAAGCTAGTATTATAATACTAATAACCTAGAT	720
Dp	661	CGTAATAATTTAGTGGATCTATACAAATAAGCTAGTATTATAATACTAATAACCTAGAT	720
Qy	721	CCACTAAATGGGGGAAAGCTTTTAATCTATAGATTACTATACACGTTAATGGGAATAT	780
Dp	721	CCACTAAATGGGGGAAAGCTTTTAATCTATAGATTACTATACACGTTAATGGGAATAT	780
Qy	781	AATAATACGTTATCAACTATTATATGAACAAAGACTTAATGCTGATGCTATATTAATGT	840
Dp	781	AATAATACGTTATCAACTATTATATGAACAAAGACTTAATGCTGATGCTATATTAATGT	840
Qy	841	TTTATTTAAAAAGTGATTTCAAAATATATGAACAAAGTTTGTGAGGACCTTTTCAAAACGCT	900
Dp	841	TTTATTTAAAAAGTGATTTCAAAATATATGAACAAAGTTTGTGAGGACCTTTTCAAAACGCT	900
Qy	901	AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTAGTGCATGATGTAAACACCGTCAT	960
Dp	901	AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTAGTGCATGATGTAAACACCGTCAT	960
Qy	961	TATAATATGCAAGAAAGACGTTTGAATGCTGATGAACCTTCAAGTGAATTC 1015	
Dp	961	TATAATATGCAAGAAAGACGTTTGAATGCTGATGAACCTTCAAGTGAATTC 1015	

RESULT 9				
LOCUS	865869			
DEFINITION	TM-1=23 kDa polypeptide	1015 bp	DNA	linear
ACCESSION	865869			
VERSION	865869.1			
KEYWORDS	GI:425376			
SOURCE				
ORGANISM	Mycoplasma gallisepticum			
REFERENCE	Bacteria; Firmicutes; Mollicutes; Mycoplasmatataceae; Mycoplasma.			
	1 (bases 1 to 1015)			

FEATURES	SOURCE
<p><b>AUTHORS</b> Saito, S., Fujisawa, A., Ohkawa, S., Nishimura, N., Abe, T., Kodama, K., Kamogawa, K., Aoyama, S., Iritani, Y., and Hayashi, Y.</p> <p><b>TITLE</b> Cloning and DNA sequence of a 29 kilodalton polypeptide gene of Mycoplasma gallisepticum as a possible protective antigen</p> <p><b>JOURNAL</b> Vaccine 11 (10), 1061-1066 (1993)</p> <p><b>MEDLINE</b> 94025893</p> <p><b>PUBMED</b> 8212828</p> <p><b>REMARK</b> GenBank staff at the National Library of Medicine created this entry [NCBI gisdbseq 138145] from the original journal article. This sequence comes from Fig. 2.</p>	<p>1..1015 Location/Qualifiers</p> <p>/organism="Mycoplasma gallisepticum" /mol_type="genomic DNA" /db_xref="taxon:2096" 202..&gt;1015 /gene="TM-1" 202..&gt;1015 /note="29 kDa polypeptide; This sequence comes from Fig. 2; conceptual translation presented here differs from translation in publication" /codon_start=1 /transl_table=4 /product="TM-1" /protein_id="AAB28343.2" /db_xref="GI:7249262" /translation="MNKKRIILKTIISLGTSPFISGISSCMSGTSTKKDANPNNGOTOLQARMEITDLINAKRTIASLODYAKIEBALSSVSAETVNNNLNTLEOLKQAKTNIRLSAINQNTDPTTPEHNEHLYEAYALKTTELEQPTNIEGLSTAYNOIRNNLIDLNNASSLITITLDPPLNGSMLLBSNIEITVRNINNTLISTNEQKTNADALSNSIKKYIÖNNSEFVETFTFNANVQPSNYSFVAFSAVDTPVNYKARRTWNQDEPSSRI"</p>
<p><b>BASE COUNT</b> 401 a 161 c 138 g 315 t</p> <p><b>ORIGIN</b></p>	
<p>Query Match 75.5%; Score 986.2; DB 1; Length 1015; Best Local Similarity 98.2%; Pred. No. 5.1e-134; Matches 997; Conservative 0; Mismatches 18; Indels 0; Gaps 0;</p>	
<p>QY 1 AAAAACTCGATGTGTAATCTGATATCTTTGGCTTAAAAAACAACAATCTTCTAACAA 60</p> <p>DB 1 AAAAACTCGATGTGTAATCTGATATCTTTGGCTTAAAAAACAACAATCTTCTAACAA 60</p> <p>QY 61 AATCTTAATTAATTAAGCCGTTAAATTAAGTAAATTAATTAATTAATTAATTAATC 120</p> <p>DB 61 AATCTTAATTAATTAAGCCGTTAAATTAAGTAAATTAATTAATTAATTAATTAATC 120</p> <p>QY 121 AACCAAAATCTCAGTAATTAAGCGTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 180</p> <p>DB 121 AACCAAAATCTCAGTAATTAAGCGTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 180</p> <p>QY 181 ATAAATATATCTTATATATTTCTATGAATTAAGAAAAAGATCATCTTAAGAAGCTATAGTTTG 240</p> <p>DB 181 ATAAATATATCTTATATATTTCTATGAATTAAGAAAAAGATCATCTTAAGAAGCTATAGTTTG 240</p> <p>QY 241 TTAGGTACAACATCTTTCTTGAAGATTTGGAATTTCTAGCTGATGTATTTACTTAAAAAA 300</p> <p>DB 241 TTAGGTACAACATCTTTCTTGAAGATTTGGAATTTCTAGCTGATGTATTTACTTAAAAAA 300</p> <p>QY 301 GATGCAAAACCAATTAATGSCCAAAACCCATTTAGACAGCGGCAATGAGTTTAAAGAT 360</p> <p>DB 301 GATGCAAAACCAATTAATGSCCAAAACCCATTTAGACAGCGGCAATGAGTTTAAAGAT 360</p> <p>QY 361 GATGCAAAACCAATTAATGSCCAAAACCCATTTAGACAGCGGCAATGAGTTTAAAGAT 420</p> <p>DB 361 GATGCAAAACCAATTAATGSCCAAAACCCATTTAGACAGCGGCAATGAGTTTAAAGAT 420</p> <p>QY 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTTAAACAATACTTAAATGCAACTTA 480</p> <p>DB 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTTAAACAATACTTAAATGCAACTTA 480</p> <p>QY 481 GAACACTAAAAATGGCTAAAACTTAATTTGAATCAGCCATCAACCAAGCTAATAGCGAT 540</p> <p>DB 481 GAACACTAAAAATGGCTAAAACTTAATTTGAATCAGCCATCAACCAAGCTAATAGCGAT 540</p>	

Db 481 GAACAATAAAATGGCTAAATCTAATTAGATCAGCCATCAACCAAGCTAATACGAT 540  
 Qy 541 AAAACGACTTTTGGATATGAGACACCAATATTTAGTTGAAGCATCAAAAGCACTAAAAAC 600  
 Db 541 AAAACGACTTTTGGATATGAGACATCAATATTTAGTTGAAGCATCAAAAGCACTAAAAAC 600  
 Qy 601 ACTTTAGAACCAACGCTGCTACTAACCCTTGAAGTTTGTCACTCACTGCTTAATCAAT 660  
 Db 601 ACTTTAGAACCAACGCTGCTACTAACCCTTGAAGTTTGTCACTCACTGCTTAATCAAT 660  
 Qy 661 CGCATATATTTAGTATCTATACATTAAGCTACTATTTAATTAATCACTAAACCTAGAT 720  
 Db 661 CGCATATATTTAGTATCTATACATTAAGCTACTATTTAATTAATCACTAAACCTAGAT 720  
 Qy 721 CCACTAAATGGGGAGCGCTTTTGAATTTCTAATGAGATTAATCACTAATCGAATAT 780  
 Db 721 CCACTAAATGGGGAGCGCTTTTGAATTTCTAATGAGATTAATCACTAATCGAATAT 780  
 Qy 781 AATAATACGTTATCACTATTAATGACAAAGACTAATGCTGATGCAATATCTAATAGT 840  
 Db 781 AATAATACGTTATCACTATTAATGACAAAGACTAATGCTGATGCAATATCTAATAGT 840  
 Qy 841 TTTATTAATAAAGTATTTCAATAATATGACAAAGTTTGTAGGGAGCTTTTCAAAAGCT 900  
 Db 841 TTTATTAATAAAGTATTTCAATAATATGACAAAGTTTGTAGGGAGCTTTTCAAAAGCT 900  
 Qy 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAAGTGTATGATCAACCCGTCAT 960  
 Db 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAAGTGTATGATCAACCCGTCAT 960  
 Qy 961 TATTAATATGACAAAGACCGTTTGAATGATGATGATCACTTCAAGTGAATTC 1015  
 Db 961 TATTAATATGACAAAGACCGTTTGAATGATGATGATCACTTCAAGTGAATTC 1015

RESULT 10  
 AX65187 1082 bp DNA linear PAT 26-MAR-2003  
 LOCUS AX65187  
 DEFINITION Sequence 24 from Patent EP1275716.  
 ACCESSION AX65187  
 VERSION AX65187.1 GI:29290312  
 KEYWORDS

SOURCE Mycoplasma gallisepticum  
 ORGANISM Mycoplasma gallisepticum  
 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1  
 AUTHORS Okuda, T., Saito, S., Dorsey, K.M. and Tsuzaki, Y.  
 TITLE Modified dna molecule, recombinant containing the same thing, and  
 uses thereof  
 JOURNAL Patent: JP 1275716-A 24 15-JAN-2003;  
 Location/Qualifiers

FEATURES  
 source  
 1.1082  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:2096"  
 /note="Modified Tm-1 portion (downstream of Bgl1) of  
 pNZ40K-S"

BASE COUNT 403 a 184 c 183 g 312 t  
 ORIGIN

Query Match 70.6%; Score 921.6; DB 6; Length 1082;  
 Best Local Similarity 98.0%; Pred. No. 1.2e-124;  
 Matches 933; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 355 ACAGATCAATCAATGCTTAAGGATGACATTAAGCTTCACTACAGCATATGCAAGATT 414  
 Db 1 ACAGATCAATCAATGCTTAAGGATGACATTAAGCTTCACTACAGCATATGCAAGATT 60  
 Qy 415 GAAGCTAGTTATCATCTGCTTATATGTAAGCTGAACAGTTAACTAATCCTTAATGCA 474  
 Db 61 GAAGCTAGTTATCATCTGCTTATATGTAAGCTGAACAGTTAACTAATCCTTAATGCA 120

Qy 475 ACATAGAACCAATAAAATGGCTAAATCTAATTAGATCAGCCATCAACCAAGCTAAT 534  
 Db 121 ACATAGAACCAATAAAATGGCTAAATCTAATTAGATCAGCCATCAACCAAGCTAAT 180  
 Qy 535 ACGGATTAACGACCTTTTGTATTAATGACACCCAAATTTAGTTGAAGCATCAAAAGCACTA 594  
 Db 181 ACGGATTAACGACCTTTTGTATTAATGACACCCAAATTTAGTTGAAGCATCAAAAGCACTA 240  
 Qy 595 AAAACCACTTTAGAACACGCTGCTACTAACCCTTGAAGTTTGTCACTCACTGCTTAAT 654  
 Db 241 AAAACCACTTTAGAACACGCTGCTACTAACCCTTGAAGTTTGTCACTCACTGCTTAAT 300  
 Qy 655 CAAATTCGATATTTAGTATGATCTATACATTAAGTAACTAGTATTAATTAATCAAAACA 714  
 Db 301 CAAATTCGATATTTAGTATGATCTATACATTAAGTAACTAGTATTAATTAATCAAAACA 360  
 Qy 715 CTAGATCCATAAATGGGGAGCGCTTTTGAATTTCTAATGATTAATCACTAATCGG 774  
 Db 361 CTAGATCCATAAATGGGGAGCGCTTTTGAATTTCTAATGATTAATCACTAATCGG 420  
 Qy 775 AATTTAATTAATACGTTATCACTAATTAATGAACAAAGACTAATGCTGATTAATCT 834  
 Db 421 AATTTAATTAATACGTTATCACTAATTAATGAACAAAGACTAATGCTGATTAATCT 480  
 Qy 835 AATGTTTATTAATAAAGTATTCATAATATGAACAAAGTTTGTAGGAGCTTTTACA 894  
 Db 481 AATGTTTATTAATAAAGTATTCATAATATGAACAAAGTTTGTAGGAGCTTTTACA 540  
 Qy 895 AAGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAAGTGTATGATCAACCC 954  
 Db 541 AAGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAAGTGTATGATCAACCC 600  
 Qy 955 GTCAATTAATAATGACAAAGACCGTTTGAATGATGATGATCACTTCAAGTGAAT 1014  
 Db 601 GTCAATTAATAATGACAAAGACCGTTTGAATGATGATGATCACTTCAAGTGAAT 660  
 Qy 1015 CTTCGAAACAGATATGATCAAGATGTTCTTGAATTAATAGTTAAGTCTGCAACAAAC 1074  
 Db 661 CTTCGAAACAGATATGATCAAGATGTTCTTGAATTAATAGTTAAGTCTGCAACAAAC 720  
 Qy 1075 ACGAATGACCAATTTAGTTTGAACATTAATGCTGATGATCACTGTTTATTAATTCCT 1134  
 Db 721 ACGAATGACCAATTTAGTTTGAACATTAATGCTGATGATCACTGTTTATTAATTCCT 780  
 Qy 1135 TATTAAGTTTAAAGACGCTGATGATTAATACCTGATTAATCAATTAATTAAT 1194  
 Db 781 TATTAAGTTTAAAGACGCTGATGATTAATACCTGATTAATCAATTAATTAAT 840  
 Qy 1195 GGAATGTTCAACAGTTGATGATGCTCACTTAATACCTGATTAATCAATTAATTAAT 1254  
 Db 841 GGAATGTTCAACAGTTGATGATGCTCACTTAATACCTGATTAATCAATTAATTAAT 900  
 Qy 1255 CCAACTCCAGAGTTGATGATTAAGTCTCAAAATCGTTTATCAGTT 1306  
 Db 901 CCAACTCCAGAGTTGATGATTAAGTCTCAAAATCGTTTATCAGTT 952

RESULT 11  
 E02348 853 bp DNA linear PAT 29-SEP-1997  
 LOCUS E02348  
 DEFINITION DNA sequence coding for TMG-1.  
 ACCESSION E02348  
 VERSION E02348.1 GI:2170583  
 KEYWORDS JP 1990111795-A/7.  
 SOURCE Mycoplasma gallisepticum  
 ORGANISM Mycoplasma gallisepticum  
 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1 (bases 1 to 853)  
 AUTHORS Kodama, K., Saito, S., Yanagida, N., Kamogawa, K., Iritani, K. and  
 Aoyama, S.  
 TITLE FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE  
 THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME  
 JOURNAL Patent: JP 1990111795-A 7 24-APR-1990;

COMMENT NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1990111795-A/7  
PD 24-APR-1990  
PF 02-JUN-1989 JP 1989140283  
PR 02-JUN-1988 JP 88P 136343  
PI KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,  
PI IIRITANI KOICHI, AOYAMA SHIGEMI  
PC C07K13/00, C07K7/06, C07K15/04, C12N1/21, C12N15/31//A61K37/02, PC  
C12P21/02, PC  
G01N3/569, (C12N1/21, C12R1:19), (C12N15/31, C12R1:35), (C12P21/02, PC  
C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: strain=56;  
FH key Location/Qualifiers  
FT 5'UTR 1..39  
FT CDS 40..825  
FT FT /product="MG-1",  
FT 826..853.  
FEATURES  
source 1..853  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
BASE COUNT 329 a 138 c 128 g 258 t  
ORIGIN  
Query Match 63.1%; Score 824.2; DB 6; Length 853;  
Best Local Similarity 97.9%; Pred. No. 1.8e-110; Indels 0; Gaps 0;  
Matches 835; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
QY 163 TTAGTCATCTTTTAAAGATATAATATATCTTAATATCTATGATATGAAAAAGATATC 222  
DB 1 TTAGTCATCTTTTAAAGATATAATATATCTTAATATCTATGATATGAAAAAGATATC 60  
QY 223 TTTAAGATATTTAGTTTGTAGTCAACATCTCTTTCTTAGCATTTGGATTTCTAGCTGT 282  
DB 61 TTTAAGATATTTAGTTTGTAGTCAACATCTCTTTCTTAGCATTTGGATTTCTAGCTGT 120  
QY 283 ATGTCTATTTACTTAAAGATGCAAAACCAATATATGSCAAACCAATTTGAAGCAGC 342  
DB 121 ATGTCTATTTACTTAAAGATGCAAAACCAATATATGSCAAACCAATTTGAAGCAGC 180  
QY 343 CGAATGAGTTAAACAGATCTATCAATGCTAAAGCAGTATGCTTCACTACAGAC 402  
DB 181 CGAATGAGTTAAACAGATCTATCAATGCTAAAGCAGTATGCTTCACTACAGAC 240  
QY 403 TATGCCAAGATTGAAGCTTATCATCTGCTTATAGTGAAGTGAACAGTTAACAT 462  
DB 241 TATGCCAAGATTGAAGCTTATCATCTGCTTATAGTGAAGTGAACAGTTAACAT 300  
QY 463 AACCTTATGCAACCTTAGAACAACTAAATAATGCTAAATTTAGATCGCCATC 522  
DB 301 AACCTTATGCAACCTTAGAACAACTAAATAATGCTAAATTTAGATCGCCATC 360  
QY 523 AACCAAGCTAATACGATTAACGATTTTGTATATGAAACCAACCAATTTAGTGAACA 582  
DB 361 AACCAAGCTAATACGATTAACGATTTTGTATATGAAACCAACCAATTTAGTGAACA 420  
QY 583 TACCAAGCACTTAAACCACTTTAGAACAGTGTCTACTAACCTTTGAAGTTTGTATCA 642  
DB 421 TACCAAGCACTTAAACCACTTTAGAACAGTGTCTACTAACCTTTGAAGTTTGTATCA 480  
QY 643 ACTGCTTATTAATCAATTCGCAATATTTAGTGAAGTCTATCAATTAAGCTAGTGTGA 702  
DB 481 ACTGCTTATTAATCAATTCGCAATATTTAGTGAAGTCTATCAATTAAGCTAGTGTGA 540  
QY 703 ATTAATAAACACTAGATCACTAAATGGGGGAACGCTTTTATGATTTCTAATGAGATTACT 762

DB 541 ATAACTAAACACATAGATCCACTAAATGGGGGAATGCTTTAGATTTCTAATGAGATTACT 600  
QY 763 ACAGTTAATGGAATTTAATAATACGTTATCACTTTAATGAAACAAAGACTTAAGCT 822  
DB 601 ACAGTTAATGGAATTTAATAATACGTTATCACTTTAATGAAACAAAGACTTAAGCT 660  
QY 823 GATCATTAATCTAATAGTTTATTTAATAAAGTGAATTCAAATAATGAAACAAAGTTTGTGA 882  
DB 661 GATCATTAATCTAATAGTTTATTTAATAAAGTGAATTCAAATAATGAAACAAAGTTTGTGA 720  
QY 883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTCT 942  
DB 721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTCT 780  
QY 943 GATGTAACACCCGTCATTTAATAATGCAAGAGACCCGTTGGATGATGATGAACCT 1002  
DB 781 GATGTAACACCCGTCATTTAATAATGCAAGAGACCCGTTGGATGATGATGAACCT 840  
QY 1003 TCAAGTAGAATTC 1015  
DB 841 TCAAGTAGAATTC 853  
RESULT 12  
E02342 708 bp DNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION DNA sequence coding for MG-1.  
ACCESSION E02342  
VERSION E02342.1 GI:2170577  
KEYWORDS JP 1990111795-A/1.  
SOURCE Mycoplasma gallisepticum  
ORGANISM Mycoplasma gallisepticum  
REFERENCE 1 (bases 1 to 708)  
Kodama, K., Saito, S., Yanagida, N., Kamogawa, K., Iritani, K. and  
Aoyama, S.  
FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE  
THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME  
Patent: JP 1990111795-A 1 24-APR-1990;  
NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1990111795-A/1  
PD 24-APR-1990  
PF 02-JUN-1989 JP 1989140283  
PR 02-JUN-1988 JP 88P 136343  
PI KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,  
PI IIRITANI KOICHI, AOYAMA SHIGEMI  
PC C07K13/00, C07K7/06, C07K15/04, C12N1/21, C12N15/31//A61K37/02, PC  
C12P21/02, PC  
G01N3/569, (C12N1/21, C12R1:19), (C12N15/31, C12R1:35), (C12P21/02, PC  
C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC \*source: strain=56;  
FH key Location/Qualifiers  
FT 5'UTR 1..708  
FT CDS 1..708  
FT FT /product="MG-1",  
FT 709..762.  
FEATURES  
source 1..708  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
BASE COUNT 278 a 119 c 107 g 204 t  
ORIGIN  
Query Match 52.1%; Score 679.8; DB 6; Length 708;  
Best Local Similarity 97.6%; Pred. No. 1.7e-89;  
Matches 690; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 280 TGTATGCTATTTACTAAAAAAGATGCAAAACCAATTAATGCGCAACCAATTAAGAACCA 339







Db 258967 TTTCAAACTTGAACCTGAAAAAACAATGCTTAACCTTGAAGCTTGAAGCTTGAAG 259026  
 QY 851 AAGTATTCMAAATATGAAACAAAGTTTGGAGCTTTTACAAACCTAATGTTCAAC 910  
 Db 259027 AGTTTATGTTAAATAATTAACAGAAATTTGATAC--AACTAATATGAGAAACAC 259083  
 QY 911 CTTCACATCAGCTTTGTTGCTTTTATGCTGATGATTAACCCCGTCAATTAT----- 963  
 Db 259084 CTGGAATATACAGCTTTGTTGTTATGTTTAAAGCAAAATATATGAGATTCCTA 259143  
 QY 964 -----AAATATCAGAAAGACCGTTTGGATGATGATGAACCTTCAAGTAAATTTCTTG 1018  
 Db 259144 ACTGGAATTTGCACAAAGAAAGGTTTGAACCTAGATGAACCGTGAACCTAGTCAATTT 259203  
 QY 1019 CAACACGAAATGTA-----TCACGATGTTTCTTGAATTAATGTTAGCTG 1066  
 Db 259204 CTAGCCTTCTGATTAATTTCTTCTACCTTAACAGAAATATGAAATTTATGTTTATCAG 259263  
 QY 1067 GAACAAACAGAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1126  
 Db 259264 GAGCAGCTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 259323  
 QY 1127 ATTTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1186  
 Db 259324 ATTTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 259383  
 QY 1187 TAAATTAATGAATTTTCAACAAAGTTGATGATGATGATGATGATGATGATGATGATGAT 1246  
 Db 259384 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259443  
 QY 1247 CAGCTAATCCACTCCAGCAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1287  
 Db 259444 CAGATTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259484

RESULT 14  
 AF210770 4568 bp DNA linear BCT 02-FEB-2000  
 LOCUS Mycoplasma gallisepticum pmga-like protein 9.1 gene, partial cds;  
 DEFINITION pmga-like protein 9.2 gene, complete cds; and pmga-like protein 9.3  
 gene, partial cds.  
 ACCESSION AF210770.1 GI:6851355  
 VERSION AF210770  
 KEYWORDS Mycoplasma gallisepticum  
 SOURCE Mycoplasma gallisepticum  
 ORGANISM Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1 (bases 1 to 4568)  
 AUTHORS Phair,G.T., Branton,S.L., Hanson,L.A., Minton,F.C., Lott,B.D.,  
 May,J.D. and Hughtalc,M.B.  
 TITLE A novel pmga-like gene from the F-strain (vaccine strain) of  
 Mycoplasma gallisepticum  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 4568)  
 AUTHORS Phair,G.T., Branton,S.L., Hanson,L.A., Minton,F.C., Lott,B.D.,  
 May,J.D. and Hughtalc,M.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-DEC-1999) College of Veterinary Medicine, Mississippi  
 State University, Box 9825, MS 39762, USA  
 FEATURES  
 SOURCE  
 1..4568  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /strain="F"  
 /db\_xref="taxon:2096"  
 <1..637  
 /note="adhesin"  
 /codon\_start=2  
 /transl\_table=4  
 /product="pmga-like protein 9.1"  
 /protein\_id="AAF29523.1"  
 /db\_xref="GI:6851356"

repeat\_region  
 CDS  
 /translation="VYNDIFSNILDKODEPNSVTYNLLKGYSLADHSFYOFSSN  
 RNESNPYLVGFIGHGRNNLNSNTNNVAPSVOTSRNLTITLVNAPKQCY  
 IKGSLITSNRNLRKPTTATANNSTITFVKGNNMSITGTPRANNIDERSGSSSG  
 QNESKTIKLNKLNKLVITTSVMNDKNKPGADYIGNLFTLLMSPTEMAKK"  
 787..831  
 /note="microsatellite"  
 /rpt\_type=tandem  
 /rpt\_unit=gaa  
 /rpt\_count=2982  
 /note="adhesin"  
 /codon\_start=1  
 /transl\_table=4  
 /product="pmga-like protein 9.2"  
 /protein\_id="AAF29524.1"  
 /db\_xref="GI:6851357"  
 /translation="MKRNILKRVSLIGISFVMLAAASCTATPTPEPPKPPMP  
 NPSPGMDNGDTPNDGMEVNSAQQLAAKELSDILATOSNLAKVADYNTONTL  
 TAAVTTASTSDNTSVLEQVKSATSTLDAIDTAASKTSFDEKPELITKAYALKE  
 TLKNEEVLISGLTDSNPAITKIMLTALYOSGDFVATLNDPVGNAPOIADITKADK  
 IADAVSKLETWKTNNATLATSFEVYEVNKLGLIDTNNRQEPNYSFVGSVANTN  
 MEIIPNNRPAQRVMTSDNGRTSLSTSDNSSTLFEVMSISLGSAGTYSLTPEY  
 GSTGYLFFPKLVKSGDENVALQYTLNLSGAOEVPAPYKTSVSAUSGDSNOT  
 ESAEIMPVTSIDNPATVSDINIAKLTLNLRKFSNTIEFSVPEPSKVAPEMGM  
 YLTSNTANEAKVNSIFGVNDNSEASTVYVDILKGYSLATWSTVTRFANUTISM  
 PENATYLVGFIGQLARTVGSIPRRNPEIPNNENRFTLVNAPKADYHISGSY  
 LTRNGLKLTVDITDKNNSITITSGKNNMTLGHFPTSKNNSNGDSVENNKA  
 SILKREGAKTVIAGTGGKNAFYIGNLFTLNSTNVAQSSST"  
 3147..3182  
 /note="microsatellite"  
 /rpt\_type=tandem  
 /rpt\_unit=gaa  
 /rpt\_count=2982  
 /note="adhesin"  
 /codon\_start=1  
 /transl\_table=4  
 /product="pmga-like protein 9.3"  
 /protein\_id="AAF29525.1"  
 /db\_xref="GI:6851358"  
 /translation="MKRNILKRVSLIGISFVMLAAASCTAATPTPEPPNPEP  
 KPPMPNPPSGMNGDTPNDGMEVNSAQQLAAKELSDILATOSNLAKVADYNTONTL  
 AKIONTILAAVTTASTSDNTSVLEQVKSATSTLDAIDTAASKTSFDEKPELITKAYALKE  
 SYKLMATLAKETITVTLNDPRKSLATLNDINGSCGEELVGHITLDPGIVYANAT  
 ITBEITKIEVISEKTIQDQKNNADQDNTQSPFLDKTLENVEDAKMOPGIVANSY  
 GYSVDVGTSGQETIIPNNRPAQRVMTSDNGRTSLSTSDNSSTLFEVMSISLGSAGTYSLTPEY  
 SLAGTGAKYTLFTFYGPSTGMLFEPYKLVKANDVDGLQYKLNSTNLTPLIFEGTT  
 TNPPAATVENINVAKRVLTGLAFKNT"

BASE COUNT 1726 a 706 c 760 g 1376 t  
 ORIGIN  
 Query Match 19.1%; Score 249; DB 1; Length 4568;  
 Best Local Similarity 55.2%; Pred. No. 3.3e-27;  
 Matches 586; Conservative 0; Mismatches 445; Indels 30; Gaps 4;

QY 251 CATCTTTCTTACGATTGGAGTTTCTACCTGATGCTTACTTAAATAAGTCAACC 310  
 Db 1097 CAACCCCTCTAGTGTATATGAATGCGGAGATCTAATCCAGAAATACGAGGAA 1156  
 QY 311 CAATATATGCGCAACCAATTAAGAGCGCGGAGTGAAGATTAACATCAATG 370  
 Db 1157 TGAAGATTTCTGCAACATTTAGCAGCTGCTAAATAAGATTAAGCATTTATTAGCA 1216  
 QY 371 CTAAAGCATGACATTAGCTTCACTAACAAGCTATGCCAAGTTGAAGCTAGTTATCAT 430  
 Db 1217 CACAAAGCAGTAACCTTGCAAAATATGCAAGATTAACAAATTTGCAATCTTAACAG 1276  
 QY 431 CTGCTTATAGTAAGCTGAAACAGTTAAACAATTAATTAAGCAACATTAGAACACTAA 490  
 Db 1277 CTGCATACCAACTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1336  
 QY 491 AATAGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 550  
 Db 1337 AAGGCTCATCACTTACAGAGAGCTATTAATGATCTGCAAGTTCAAAACTAGTT 1396



```

Oy 551 TTGATATGACACCCCAATTGTTGAGACATACAAAGCACTAAAAACACTTTAGAAC 610
Db 1397 TTGATGAAAAAATCTGATATATCAAGCATATTATGCTTTAAAGAAACATTAATAA 1456
Oy 611 AACGCTACTACCTTGAAGTTTGTGATCATCACTGCTTATATCAAAATTCGCATATT 670
Db 1457 ATGAAGAAACAGTTTAAAGTGTAACTGATTTCAATTTTGCAACAAATCAAACTAAC 1516
Oy 671 TAGTGATCTATACAAATTAAGCTAGTATGTTTAAATTAACAAACATAGATCACTAAATG 730
Db 1517 TTACGCTTTATATATCAATCTGTTAAAGATTTTGTTTAAAGCAACATTAAGTCTGT- 1573
Oy 731 GGGGAAACCTTTAGATTCTTAATGAGATTACTACAGTAATCGAATAATTAATTAATACGT 790
Db 1574 GTGAAAGAGCTCCCAATTCGTGATTAACAAGACAGATTAAGATATGCTGATGCTG 1633
Oy 791 TATCAACTATTAATGACAAAAGACTAATGCTGATGATATCTAATAGTTTATTAATAA 850
Db 1634 TTTCAAACTTGAAACCTGAAAAACAAATGCTAATTAATTAAGTTTGTGTTAAG 1693
Oy 851 AAGTATTCAAAATTAATGACAAAGTTTGTAGGAGCTTTTCAACGCTAATGTTCAAC 910
Db 1694 AGGTTTATGTTAAATTAATTAACAGGAATGATAC--AACTAATTAATGAGAACACAC 1750
Oy 911 CTCAACTACAGTTTGTGCTTTTGTAGTGTATGATGTAACACCCGCTCAATTAAT- 963
Db 1751 CTGAAATTTACAGCTTTGTTGTTAGTGTAAATGCAACAAATTAATGATGATTCCTA 1810
Oy 964 -----AATATGCAAGAAAGACCGTTTGAATGATGATGAACTTCAAGTAAATTTCTG 1018
Db 1811 ACTGGAATTTGCAACAAAGAAAGTTTGAACATGATGAAACGTAAGTAACTTAATTT 1870
Oy 1019 CAACACGAATATGTA-----TCACAGATGTTCTTGAATTAATAGTTTAGCTG 1066
Db 1871 CTAGACATCTTGATTAATCTTCTACCTTAACAGAGTATCATGAAATTTATGATTAACG 1930
Oy 1067 GAACAAACAGAAATGACAAATTTAGTTTAGCACTATAGTGCATCAACCTGTTATTTAT 1126
Db 1931 GAGCAGATTAATTAATTAACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1990
Oy 1127 ATTTCCCTTATTAATGTTGTTAAAGCACTGATGCTAATTAACCTTGATTAACATCAAT 1186
Db 1991 ATTTCCCATATTAAGTTAGTTAAAGAGGATGAAATTAATTAATTAATTAATTAATTAAT 2050
Oy 1187 TAAATTAATGGAATGTCACAAAGTTGATGCTTCCACTTCAACTGACGCAATTAATACGA 1246
Db 2051 TAAATTAATGATGATGACAAAGTTAATTTTGTCTCCAACTGTAATAAATAGTATG 2110
Oy 1247 CAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCT 1287
Db 2111 CAGATTCTTCTGTGATTCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2151

```

RESULT 15  
AE016969/c 301903 bp DNA linear BCT 09-JUN-2003

DEFINITION Mycoplasma gallisepticum strain R section 3 of 4 of the complete genome.

ACCESSION AE016969  
VERSION AE016969.1  
KEYWORDS GI:31541483

SOURCE Mycoplasma gallisepticum R  
ORGANISM Mycoplasma gallisepticum R

REFERENCE 1 (bases 1 to 301903)  
Gearty,S.J., Papadistil,L., Kutish,G., Gorton,T.S., Mahairas,G., Swartzell,S., Madan,A., Nguyen,D.K., Markham,P., Browning,G., Kamal,M. and Liao,X.

TITLE The complete genome sequence of the avian pathogen Mycoplasma gallisepticum strain R

JOURNAL Microbiology (2003) In press

REFERENCE 2 (bases 1 to 301903)  
Gearty,S.J., Papadistil,L., Kutish,G., Mahairas,G., Swartzell,S.,

# TITLE JOURNAL

Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G., Mustafa,K. and Liao,X.  
Submitted (17-OCT-2002) Department of Pathobiology and Veterinary Sciences, and Center of Excellence for Vaccine Research, The University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT 06269-3089, USA

## FEATURES source

1..301903  
/organism="Mycoplasma gallisepticum R"  
/mol\_type="genomic DNA"  
/strain="R"

/db\_xref="taxon:233150"  
complement (77..754)  
/locus\_tag="MGA\_0090"

complement (77..754)  
/locus\_tag="MGA\_0090"

/note="MGA\_0090; [R] COG1881 Phospholipid-binding protein; MGR\_456"

/codon\_start=1  
/transl\_table=4  
/product="predicted phospholipid-binding protein"

/protein\_id="AAP56784.1"

/db\_xref="GI:31541484"

/translation="MRIMKXYSIGNSKINSKVGADGYLALHGGHKNENQTPPSFSLDSWEKIPNAKSYAVIEDPAAVIGSFIMIVIKTHNIAENOSYLPHNKMLBSKOSAGENDLMQGHNSVANKTLVANKTKNDLGILPEGFTSLSENSLMYGCYPPNDHTYTVVVGGLDVDASELTHYKHNRTQKYLHNKRYVYGDPMQALHKHYVGGHTLHFKYKVG"

complement (767..1396)  
/locus\_tag="MGA\_0091"

complement (767..1396)  
/locus\_tag="MGA\_0091"

/note="MGA\_0091; [R] COG1881 Phospholipid-binding protein; MGR\_457"

/codon\_start=1  
/transl\_table=4  
/product="predicted phospholipid-binding protein"

/protein\_id="AAP56785.1"

/db\_xref="GI:31541485"

/translation="MKRTISYIFNLATYFNDYILMTIMDMKKHAIKSNCFETGDSVYLSKTVKLNKNRYESLDLTWDKIEGAKSYAVVAVYASRVIGSFIMHVNKESNKLAVYANINKSIYVQGLNSTPQCTESSKGVIIIECVGAFKNTAEASDYLPLPPDHTLYLTVRVYGLDVHDVLDPKHFSLTDLNQKTIHHTVGEHLEHFWYKEN"

complement (1437..2306)  
/gene="ara1"

/locus\_tag="MGA\_0093"

complement (1437..2306)  
/gene="ara1"

/locus\_tag="MGA\_0093"

/note="MGA\_0093; ARA1 [R] COG0656 Aldo/keto reductases related to diketoglutarate reductase; MGR\_458"

/codon\_start=1  
/transl\_table=1  
/product="ARA1"

/protein\_id="AAP56786.1"

/db\_xref="GI:31541486"

/translation="MYLTTRKKIKMKYNYLANNYKPSIGRTYKLEDKQTVAVKTALEVGTHLDCAEIEYENQSVSGIKLSGVNRKIFITSKIANDDKGYESKKAFFHKILKLDLLEYLDLLHWPILGKGFKNWQVNAQTWAKMAEFLERKAIKIGSNFLVHIEALKSKAKILPMVNOLEFHFGYLOPEIVECCQKNINIVQAMSPFSSKYLENTLVOLAKKYEVTVADQILINWIIQKDIYVLPKSKTPERIRSNFVDVNFQINDLKVINDLCCGGLHPDVDP"

2541..2930  
/locus\_tag="MGA\_0094"

2541..2930  
/locus\_tag="MGA\_0094"

/note="MGA\_0094; MGR\_459"

/codon\_start=1  
/transl\_table=4  
/product="conserved hypothetical"

/protein\_id="AAP56787.1"

/db\_xref="GI:31541487"

Query Match	Similarity	Score	DB 1	Length	301903
Beet Local	56.2%	Pred. No. 5.3e-24			
Matches	516	Conservative	0	Mismatches	372
				Indels	30
				Gaps	4
gene					
CDS					
product="PepC"					
/protein_id="AAP56792.1"					
/db_xref="GI:31541492"					
/translation="MKSNNKNNKIDPQYILKRLSLYOKREYLLTRNALVANDINLM					
/AAVPEADKNTLEFLVDAHKPTLITINQDSGRCTWFGMLPLRQTAKEKLSFV					
/SQYVFMWRYERANVFLINKIIEKADVELDRDLKELQSGDQSGWGFENLVKY					
/GLVPEVMPDSFGSHNTFILLNELQVLIKATKEIRAHKKSQKQEVADLTLLKYLE					
/MLVIAVGPSPKQPMQVADAKDQENKEELANKEKSTGTATYKTPAQOEEKAKQ					
/LKEIKTEPFSIKQITPLEPFGKOVNOYDELDLMTGTTIDYKINQKSLKDSNLF					
/EADFNPLANDRNLKRPALANLYAKQTPMFAQVNNRNNKGTGDFNDQYOSLNF					
/IDSVENDKQVRSFISNHAHTLSGVDPAKSLIKQELVKYTKNLKPKDQYQFVL					
/DLQTFEPFKKLIENSGEYKNGFVYMDQWMDNDYIDIVISTKADDFEKNRFTI					
/DPLDIGLQKLTKEELIKAGLDTPILIDGDPUGANNKGVK"					
complement (6968..7402)					
/gene="smpB"					
/locus_tag="MGA_0102"					
complement (6968..7402)					
/gene="smpB"					
/locus_tag="MGA_0102"					
/note="MGA_0102; smpB [O] COG0691 tmRNA-binding protein;					
MGR_465"					
/codon_start=1					
/transl_table=4					
/product="SmpB"					
/protein_id="AAP56793.1"					
/db_xref="GI:31541493"					
/translation="MMRLVNNKRAKXNYVELLDREAGISLSGNEVKSIALHGHKLD					
/SVYIRKEAVLNLILPKKFPDTSKVLNETRTKLLHKSIELIKDILIKQHSVLII					
/PYQIYFNNKIKVSIYLARPKRYDKQTIKERELNKIRY"					
complement (7399..8148)					
/locus_tag="MGA_0103"					
complement (7399..8148)					
/locus_tag="MGA_0103"					
/note="MGA_0103; MGR_466"					
/codon_start=1					
/transl_table=4					
Query Match	17.1%	Score 222.8	DB 1	Length 301903	
Beet Local	56.2%	Pred. No. 5.3e-24			
Matches	516	Conservative	0	Mismatches	372
				Indels	30
				Gaps	4
gene					
CDS					
product="PepC"					
/protein_id="AAP56792.1"					
/db_xref="GI:31541492"					
/translation="MKSNNKNNKIDPQYILKRLSLYOKREYLLTRNALVANDINLM					
/AAVPEADKNTLEFLVDAHKPTLITINQDSGRCTWFGMLPLRQTAKEKLSFV					
/SQYVFMWRYERANVFLINKIIEKADVELDRDLKELQSGDQSGWGFENLVKY					
/GLVPEVMPDSFGSHNTFILLNELQVLIKATKEIRAHKKSQKQEVADLTLLKYLE					
/MLVIAVGPSPKQPMQVADAKDQENKEELANKEKSTGTATYKTPAQOEEKAKQ					
/LKEIKTEPFSIKQITPLEPFGKOVNOYDELDLMTGTTIDYKINQKSLKDSNLF					
/EADFNPLANDRNLKRPALANLYAKQTPMFAQVNNRNNKGTGDFNDQYOSLNF					
/IDSVENDKQVRSFISNHAHTLSGVDPAKSLIKQELVKYTKNLKPKDQYQFVL					
/DLQTFEPFKKLIENSGEYKNGFVYMDQWMDNDYIDIVISTKADDFEKNRFTI					
/DPLDIGLQKLTKEELIKAGLDTPILIDGDPUGANNKGVK"					
complement (6968..7402)					
/gene="smpB"					
/locus_tag="MGA_0102"					
complement (6968..7402)					
/gene="smpB"					
/locus_tag="MGA_0102"					
/note="MGA_0102; smpB [O] COG0691 tmRNA-binding protein;					
MGR_465"					
/codon_start=1					
/transl_table=4					
Query Match	17.1%	Score 222.8	DB 1	Length 301903	
Beet Local	56.2%	Pred. No. 5.3e-24			
Matches	516	Conservative	0	Mismatches	372
				Indels	30
				Gaps	4
gene					
CDS					
product="PepC"					
/protein_id="AAP56792.1"					
/db_xref="GI:31541492"					
/translation="MKSNNKNNKIDPQYILKRLSLYOKREYLLTRNALVANDINLM					
/AAVPEADKNTLEFLVDAHKPTLITINQDSGRCTWFGMLPLRQTAKEKLSFV					
/SQYVFMWRYERANVFLINKIIEKADVELDRDLKELQSGDQSGWGFENLVKY					
/GLVPEVMPDSFGSHNTFILLNELQVLIKATKEIRAHKKSQKQEVAD					

```

Qy      803  ATGAACAAAGACTAATGCTGATGCAATTAATAGTTTATTAAAGATTCAAA 862
Db      22060 AACATGAAAAATATGCAAAACATGATCTGATAGTTTAAAGAAACCTAGATA 220001
Qy      863  ATATGAACAAAGTTTGTAGGACTTTACAAAGCTAATGTTCAACCTTCAACTACA 922
Db      22000  AAGAACAAATTAACAAGTAATGGGAATTACA-----CAACCTGCAAAATTACA 219953
Qy      923  GTTTGTGCTTTTATGCTGATGTAACACCCGTCATTTATTAATATGCAAGAGACCG 982
Db      219952 GTTTGTGCAATATATCAAGTATTAATCTTACTTATTAATTTTGTAAAGAGTAG 219893
Qy      983  TTGGAATGATGATGAACCTTCMAAGTAATTTCTTC-----AAACAAGATAGTA 1033
Db      219892 TTTGAAAACCAAGAAAGTGGTCAAGTACCTATGTCATTAAGAAATCAAGTGATT 219833
Qy      1034 TCACAGATGTTCTTGTGATTTATAGTTTACTGGAACAAACAGAGTACCAATTTAGTT 1093
Db      219832 TAACGAGATATCATGAATTTATAGTTTACTGGAATGAACTAATATAGTTTACTT 219773
Qy      1094 TTAGCACTATGATGTCATCACTGTTATTATATTTCCCTTATTAAGTTGTTAAGCAG 1153
Db      219772 TTGCTAATTAATGCAACACTACAGATATCTTATTTCCATACAAAGTCAAGTTAATCA 219713
Qy      1154 CTGATGCTAATTAACGTTGATTTACAATACAATTAAATGAATGTAAGTTTCAACAAGTTG 1213
Db      219712 GCG-----ATAATGTTGATTAACAATAAAGTTAATGATGCTAGTCCGGTATCTATCG 219659
Qy      1214 AGTTTGCCACTTCAACTA 1231
Db      219658 ATTTTAATTACTCAGCAA 219641

```

Search completed: August 25, 2003, 04:51:42  
 Job time : 5038 secs

THIS PAGE BLANK (USPTO)

```

Oy 293 YARRTWNGDEPSSRLANTNSITDVSWISLA-----GTNTKYQFSPSNGP 340
Db 724 KQKQDALKQVTSKQRYNAVNTSIQOTANELNTANGQLOHGIDDENATKOTOKTIDAEOSK 783
Oy 341 STGYLFFPKYKLVKADA-----NNVGLQVYKLN-----NGNVQVEFATSTS 381
Db 784 KTAIV-----DQAVVAAKALILNKQTSNSDKAAVDRALQOVSTSKALNGDALAEKKAAR 839
Oy 382 AN-----NTTANPFAVDEIKAKIVLSGLRFGQNTIELSVFPGEGNN-----KV 427
Db 840 QNLGLNHTITVAQRTALEGQINQATTV-----DGVNTVKTANTALDGMNSLQGAINDKD 894
Oy 428 APMIGNIYLSNEN 441
Db 895 ATRNQNYLDADES 908

RESULT 13
C89921
hypothetical protein ebhb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A69758; MUID:21311952; PMID:11418146
A:Accession: C89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <KUR>
A:CROSS-references: GB:BA000018; PID:g13701233; PIDN:BA042528.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: ebhb

Query Match 7.2%; Score 166.5; DB 2; Length 3890;
Best Local Similarity 19.7%; Pred. No. 0.96;
Matches 100; Conservative 89; Mismatches 210; Indels 109; Gaps 18;

Oy 17 YGNTSSPST-----QNTSREVSSVOLSEESTFYLCPPVGSV 57
Db 2427 HVSNSPSTLTPRAHTVNTTEIVDYGSNVTAIEINNAVQVANKRT-----ATI 2476
Oy 58 -----IRLEFGCMSITKQDANPNNGQT-----QLEAARMELTDLINAKA 96
Db 2477 KNGTAMPNTNLGGSTTTTPVTYNDGSTEVEQESIFTKADRELTITAKNHLDDPVSREG 2536
Oy 97 MTLASLQDY-----AKIEASLSAYSEAEYNNNNLNTLEQLKMAKTINLESAINQANTKT 152
Db 2537 KKPGTITTYNNMNAHQOINTAKTEAQOVINNEBATPOQVSDALTTKYRAQTKIDOKA 2596
Oy 153 TPDNEHPN--LVEAYKALKTLEQRATNLEGLSTAYNOIRNNILVDLVN-----KASSLIT 206
Db 2557 LLQNEEDNSQVLTSSKNLTQSSVNO-----VPSTA--GMTQOSIDNTNAKKRELEIT 2647
Oy 207 KTLDLPLNGQTLDSNEITTVNNINNTLSTINEQKTNADALSNSFIKVIQNNESFVGT 266
Db 2648 AAQRVIDNGD-ATAQOISDEKRRVDNALTLALNQAHDLADLADTHALEQAVQQLNR---TGT 2703
Oy 267 FTN---ANVQPSNYFVAFASADV-----PV-NYKRYARTWNGDEPSSRLA 310
Db 2704 TTGKKPASTAYANNSIRALQSDLTSAKNSANAIQKPIRTVQEVOSALTNNRVNERLTQ 2763
Oy 311 NNTNSITDVSWISLAGTNTK-----YQFSPSNYSGVLYFPYKLVKADANNVGLQYKL 366
Db 2764 AINQVLPLADNSALTKTKLDEBEINKSVTTTGMTQSSIQAYENAKRAGQITTTAQNVI 2823
Oy 367 NNGNVQVEFATSTS-----TSANNTTANPTPAVDEIKAKIVLSGLRFGQNTIELSV 418

```

```

Db 2824 NNGDATDQOIAAEKTYEKEKNSLQKQAIAGLTPDLAFLQTRAKTQL-----QNDID--QP 2875
Oy 419 TGEENMKVAPMIGNIYLSNENNAKDI 446
Db 2876 TSTTGMTSASVAAFNDKLSAARTKIOEI 2903

RESULT 14
AG2422
hypothetical protein all4935 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG2422
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-661 <KUR>
A:CROSS-references: GB:BA000019; PIDN:BA076634.1; PID:g17134073; GSPDB:GN00179
A:Experimental source: strain PCC 7120
A:Genetics:
A:Gene: all4935

Query Match 7.1%; Score 166; DB 2; Length 661;
Best Local Similarity 22.7%; Pred. No. 0.93;
Matches 108; Conservative 50; Mismatches 213; Indels 104; Gaps 18;

Oy 1 MHYFRNCIFFLI--VILYGNSS-----PSTQNTSREVSSVOLSEESTFYLCPPV 53
Db 1 MNLVTQAILGLITTAALGTGLASIKPATYITOTTISAN-VKPIQVAIKEA-----DEAI 54
Oy 54 GSTVIRLEFGCMSITKQDANPNNGQTQLEAARMELTDLI-----NAKAMTLASLQDY 105
Db 55 -----KK---AGEADALKKAEEDLSGIITSTRAKSDANSRLURAQDF 96
Oy 106 AKIEASLSAYSEAEYNNNNLNTLEQLKMAKTINLESAINQANTKTTPDNEHPNLYEAY 165
Db 97 TQADAGFQIQTALNTATNNRNNAQNALNTATNNRNNAQNALNTATNNRNT-----AQ 149
Oy 166 KALKTLEQRATNLEGLSTAYNOIRNNILVDLVNKASSLITKTLDPGLDLSNEITT 225
Db 150 NALNTATNNR--NNQNALNTATNN--RNNAQNALNTATNNRNNAQNALN-----TAT 198
Oy 226 VNR--NINNTLSTINEQKTNADALSNSFIKVIQ---NNEQSFVGTFTTNAVQPSNYSEVA 281
Db 199 NNRNNAQNALNTATNNNNNAQABVDYATRLAQARRNSQKAIQNALNALNQANTRLNTA 258
Oy 282 PSADVTPVNTKYARTWNGDEPSSRLANTNSITDVSWISLAGTNTKQFSPSNGPS 341
Db 259 NTAINTATNNENTANTAFN-----TANNNEST-----ANTAFNTATNNFSTA 300
Oy 342 TGYLYFPYKLVKADANNVGLQYKLNNGVQVEFATSTSANNTTANPTPAVDEIKAKI 401
Db 301 -----NTALNTATNNFNTAIALDQANTRKLNARN-----DPNTANS 337
Oy 402 VLSGLRFGQNTIELSVFPGEGNNMKVAPMIGNIYLSNENNAKDI PGYRRPGETL 456
Db 338 NFSRTGMEINATNTNFTNTANTFTTATNTFPN--ASSRRNTAEQARQVRETRL 390

RESULT 15
D90011
FntB protease [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

```

A:Genetic code: SGC3

7.3%; Score 170.5; DB 2; Length 4688;

Query Match Best Local Similarity 21.1%; Pred. No. 0.75; Indels 175; Gaps 25;

Matches 112; Conservative 79; Mismatches 165; Indels 175; Gaps 25;

19 TNSPSTQNTVSRVSVSVLSSESTP-----YICPPVGSYVIREFGCMST 67

1387 SNNVSLGNKNTNIEITKRIILVNDQSLISAKIEVDIDVINTADKKNIVQLE----- 1440

68 TKKDANPN-----NGOTLEAKRMELT-DLINAKAMTLASLODYAKIEASLS----- 114

1441 NSNDANNMLKLATNAQVAVENKKEFLKFDLSLKLIN-----QNYVYKISPSNKPATAYF 1495

115 -----AYSEAE-TVNNNLINAT-----LEOLKMAKTN 139

1496 NPTNKKNTNIYVSDEQKISLSNNINFTSYKPTINKDKSVNIDVDLQVQKOLLANQY 1555

140 LESAINQANTDKTT-----PDNEHPNLVEAYKALKITLBOQATNLEGLSTANQIRNN 193

1556 LRLKIKQINDNKVTWTDPLIFNN--NAKISFK-LSNLIHNRAYELEGL--YFPDQNS 1608

194 LVDLYNKASLITKLDPLANGSTLDSNEITTVNNNNINNTLTSTIN-----EOKTNAD 245

1609 VNDMTNNOISFNSKIHKP--KIEFEPSSLITINQTNNAIKTVSAHNAQVHFKLKTNDE 1664

246 ALSNSFIKKVIONNEQSFVGFPTNANVOPSNYSFVAFSADVTPVNYKYARRTVWNGDEPS 305

1665 ALENQIVYEA-----FAPTNNDQKV-----E 1689

306 SRIANTNSITDVSMTYSLAG--TNTKIQF--SFSNYGSPGTYVYFVKLVKADANNV 360

1690 AKINNVTVSNFNGEELFNLGRLREFTYKRLIKVTEKN-----PKAYELINKNGV 1740

361 GLOYKLNNGVOVEFAT-----STSNANTTANPTPAVDEI-----KVAKIV 402

1741 IREYK-NQS-QAYETTKQFHKHVIDVVSSTMTQOEIVKIDGIRANNNKKLELY 1797

403 LSGLRFGQNTIELSVPTGEGNNKVAIPMIGNTIYLSNENNAKTIQYRRPG 453

1798 YESNIIIDPEIKITV--DNNNN-----SVHLSPDKKEVNLVNLNPKG 1838

RESULT 11

JC6009 surface-located membrane protein Imp3 precursor - Mycoplasma hominis

C:Species: Mycoplasma hominis

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C/Accession: JCG6009

R:Adeloided, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.

J. Bacteriol. 178, 2775-2784, 1996

A>Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis Imp gene system

A:Reference number: JCG6009; MUID:96213016; PMID:8631664

A:Accession: JCG6009

A:Molecule type: DNA

A:Residues: 1-1302 <LAD>

A:Cross-references: EMBL:X55601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336

A:Gene: Imp3

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology

C:Keywords: duplication; membrane protein

us-09-901-572a-3.rpt

Page 5

Query Match Best Local Similarity 20.9%; Pred. No. 1.3; Indels 122; Gaps 19;

Matches 103; Conservative 77; Mismatches 192; Indels 122; Gaps 19;

67 ITKRDANPNQO-----TOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSEATV 122

864 ITKRIETFKDQVFKLEQTRKQIDEFIN-----TNKTNPYSTLISELTSKRDSKSI 919

123 NNNIANTLEQLMAKNTLESAINQANTDKTTPDN-----EHPN-----LYEAYKAL 168

920 TNSNKS--DIETATTELKQALAKANTDKQADNLTARSTKEQINKSISANTLAKLTDK 977

169 KTTLEQATNLEGLSTANQI--RNNVDLYNKASL-----ITKLDPLNGSTLDSN 221

978 DNTIQAKTELEKEVQKAAQVAVSNNTASQAKSLDAKTEITIKLETFKQDQVFR 1037

222 EITTVNRNINNTLTSTINEQTNADALNSFIKKVIONNEQSFVGFPTNANVOPSNYSFVA 281

1038 ELEQTRKQIDEFIN--NKNPNYSTLISELTSK--RDSKNSITNSNKSIEDTANTTEL-- 1092

282 PSADVTVPVNYKYARRTVWNGDEPSRIILANTNSITDVSMTYSLAGNTTYQSFSGYGS 341

1093 -----KQALAKAN--TDKAQADNLA-RSTKEQINKSISAN 1125

342 TGYLYFPYKLVKADANNVGLQYKLNNGVOVEFATSTSNANTTANPTPAVDEI--KYA 399

1126 T-----LAKLTDKQNTIQAKTE--LEKEVQKAAQVAVSNNTYVSQAKSSLDITYT 1176

400 KIVLSGRFGQNTIELSVPTGEGNNKVAIPMIGNTIYLSNENNAKTIQY 449

1177 EITKRIETFKDK-----EAKFNELEKTRQIQIEFTNTNKN-----PNY 1216

RESULT 12

B89921 hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C/Accession: B89921

R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani, O.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; I. Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: B89921

A:Molecule type: DNA

A:Residues: 1-6713 <KUR>

A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GND0149

A:Experimental source: strain N315

A:Gene: ebha

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology

C:Keywords: duplication; membrane protein

A:Genetic code: SGC3

C:Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology

C:Keywords: duplication; membrane protein

A:Genetic code: SGC3

## RESULT 7

VGBERR

glycoprotein B precursor - Marek's disease virus (strain RB18)

C:Species: Marek's disease virus

C&gt;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jun-2000

C:Accession: A32402; B32402

R:Rose, L.J.N.; Sanderson, M.; Scott, S.D.; Bime, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog

A:Reference number: A32402; MUID:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 &lt;ROS&gt;

A:Cross-references: GB:DJ13713; EMBL:DD00506; NID:9221836; PDB:BA02866.1; PID:9221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271,304-330 &lt;ROS2&gt;

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-21/Domain: signal sequence #status predicted &lt;STG&gt;

F:22-865/Product: glycoprotein B #status predicted &lt;GPB&gt;

F:709-728/Domain: transmembrane #status predicted &lt;TM1&gt;

F:732-752/Domain: transmembrane #status predicted &lt;TM2&gt;

F:27,184,332,364,406,425,631/Binding site: carbohydrate (Aan) (covalent) #status predict

Query Match

Best Local Similarity 13.5%; Score 314; DB 1; Length 865;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 MHYFRNCIFFLIVLILYGTNSPSTQNTVSREVSVQSEESTFYLCPPVSGSTVIRL 60

DB 1 MHYFRNCIFFLIVLILYGTNSPSTQNTVSREVSVQSEESTFYLCPPVSGSTVIRL 60

Query 61 E 61

DB 61 E 61

RESULT 8

major surface protein (clone pmga1.6) - Mycoplasma gallisepticum (fragment)

C:Species: Mycoplasma gallisepticum

C&gt;Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S51560; S48757

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organization of the multigene family which encodes the major cell surface P

A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S51560

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-320 &lt;MAR&gt;

A:Cross-references: EMBL:L28424; NID:9535687; PDB:AAA62415.1; PID:9535688.

C:Genetics:

A:Genetic code: SGC3

Query Match

Best Local Similarity 8.2%; Score 189.5; DB 2; Length 320;

Matches 46; Conservative 11; Mismatches 26; Indels 11; Gaps 3;

Query 354 AADANNVGLQYKLNNGNVQVEPATSTANNTTANPTPAVDKIVAKIVLSGLRFGONTI 413

DB 41 SADSNN-----PTNQSNSQSNQALPASA---MNETPTVDGIVAKIVLTLDLKFGSNTI 90

Query 414 ELSVP-TGEGNNKVAPEMIGNIYLSSENENADKI 446

DB 91 ELSVPTTDEGTSKVAPEMIGNIYLSSENENADKI 124

RESULT 9

S41539

fibrinogen-binding protein - Staphylococcus aureus

N:Alternate names: clumping factor

C:Species: Staphylococcus aureus

C&gt;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Oct-1999

C:Accession: S41539; S36630

R:McDevitt, D.; Francoise, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staph

A:Reference number: S41539; MUID:94224142; PMID:8170386

A:Accession: S41539

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 &lt;MCD&gt;

A:Cross-references: EMBL:Z18852; NID:9397525; PDB:CAA79304.1; PID:9397526

Query Match

Best Local Similarity 7.6%; Score 176.5; DB 2; Length 933;

Matches 112; Conservative 66; Mismatches 184; Indels 131; Gaps 23;

Query 2 HYFRNCIFFLIVLILYGTNSPSTQNTVSREVSVQSEESTFYLCPPVSGSTVIRL 52

DB 9 HAIRKSIQVAVLVGLIFGGLSKKADASSENSVTSQSSASNSKSSNSVSAAPKT 68

Query 53 VGSTVIRLEFGCMSITKKDANPNNGOTOL--EARMELTDLINAKAMTLASLQDYAKIEA 110

DB 69 DDTNV-----SDTKSSNVNNGFSTVAQNPQGETTQSSSTNATY----- 108

Query 111 SLSSAYSEAEVNNNNLNAITLQLMKNTLESANQANTDKTTPDNHNPVYKALKT 170

DB 109 EETVTEGATTTTNTQANTPATTSNTNABELNQ--TSNETTND--TNTVSSVNS-- 162

Query 171 TLEORATNLEGLST-----AVNQIRNNLVDLNLK-----ASSLIT 206

DB .163 --PQNSTNAEVSSTQSTTEATPSNNESAPQSDANQVNVAVNTSAPRMAAFLLA 220

Query 207 KTLDPNLNGTLLDSNEITTVNRNINNTLSTINEQ---KTN-ADALSNSPIK---KYIQ 257

DB 221 VAADAPAAAGTDI--TNQLTNVGVGIDSGTVPVPHQAGYVKLVNGFSPVNSAVKCDPTFRTV 279

Query 258 NNEQSPVGTPTNNAVQSNVSVFAFSADVPVNVKARVTVNGDESSHLANTNSITD 317

DB 280 PKEINLVGVSTAKVP-----IMAGDQ---VLANGVIDSD 312

Query 318 VSMIYSLAG--TNTKYQSFNSVYGPSTGYLPYKLVKADANNGLQYKLNNGVQVEF 376

DB 313 GNVYITFTDYNTQDDVKATITMFA--YI-----DPENV-----KKTGV----- 350

Query 377 ATSTSNNTTANPTPAVDKIVAKI-----VLGLRFGONTIELSVPTGEGNNKVAPE 429

DB 351 TLATIGSGSTANKTVLVDYERKGFYNIKIGTIDQIDKTNNTYRQTIYVNPSSGDNIAP 410

Query 430 MI-GNIYLSSEN 441

DB 411 VLTGNLKNPTDSN 423

RESULT 10

hypothetical protein U482 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C&gt;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: F82885

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: F82885

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 &lt;GIA&gt;

A:Cross-references: GB:A0002145; GB:AF222894; NID:96899476; PDB:AA30894.1; GSPDB:GN001.

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U482

Query Match	31.9%	Score 741.5	DB 2	Length 649
Best Local Similarity	40.0%	Pred. No. 3.7e-33		
Matches 1/6	Conservative	65	Mismatches 136	Indels 63
Gaps				12
QY	53	VGSIVLRLEFGGMSIIRK-----DANPNNGQ-----TGLEAARELND	90	
DB	15	IGSFVWLAASCTTPPTPTPNPNPPSGGNGGDTNPDDGGGMMNAASOELAAAREGLTT	74	
QY	91	LIAKAMTLASLADYAKIKELASISAYSEAEFTYNNNIANTLEQLQAKTINESAINOANTD	150	
DB	75	IFPSKAKNIGLYVDYIKKTOITLTKAYDAKAYLDNBSSTTQUNLEAKTRLETAIRPAATS	134	
QY	151	KTFEDNEHPNLVEAYKALKTTLEQPRATNIEGLSTAYNQIKNNTVDLYNKASLSIKRTLD	210	
DB	135	KQFDEQHAELVYKYEKELKTTLSNETATLAPYADQYAGIKHSLSGLYDAKALITKTLE	194	
QY	211	PLNGGTLIDSENETITTYNRBNINNTL--STINEQKTADLASNSFIKAYIONNEQSFWGTFT	268	
DB	195	PVBSDDP-LTASAVMMANTKI VEAIKDEVLNPOKEMANTLADSPVQYLVREKTI GVEEAH	253	
QY	269	NANVQEPNSYFVAFSADVTPV-----NYKIAARTW-NGDESSKILANT-----	312	
DB	254	N-KAQPANYSPFVGSGVDITGTTGTQTSIPNMVYAAKRIITFTSDEP--RSISINTADGQTM	310	
QY	313	-NSTIDVSWIYSLAGNTTKYQSFNSYVGSGTGLVFPYKLVLAADANNVGLQYLRNNGV	371	
DB	311	AQPLNSWISWISLAGTAKKLTLEFTYVGTGLVFPYKLVNTSPQVGLGLEVYKLD---	367	
QY	372	QOVEFATSTSA-----NNTTANPTPAVDEIKVAKIYLSGLRPGQNTIELSPTEBGMMNK	426	
DB	368	-----ATKPSAIRTEGSDQTMNGKTPTVNDINAKVTLANLNGSKLIEFSVA-----EK	417	
QY	427	VAPMIGNIYLLSSNNENADKI 446		
DB	418	VSPMIGNIYLLSSSPNNMKI 437		

```

RESULT 5
A49218
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P. F.; Glew, M. D.; Whitear, K. G.; Walker, I. D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218, NUID:93162830, PMID:8432610
A:Accession: A49218
A:Status: preliminary
A:Molecule type: DNA; Protein
A:Residues: 1-647 <NA>
A:Cross-references: GB:55216; NID:9265625; PIDN:AAB25397.1; PID:9265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIP:125183)
C:Genetics:
A:Genetic code: SGC3

Query Match      30.6%; Score 711.5; DB 2; Length 647;
Best Local Similarity 39.0%; Pred. No. 1,6e-31;
Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;

QY      53  VGSFVLRERFGCMSITKK-----DANPNNGQ-----TQLEAAMELTD 90
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      15  IGSFVMLAASCTTPNPPTPNPNPPSGGNGGDTNPGGQGMMAAASQELAAARGLTT 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      91  LINAKMTLASLQDYAKIEASLSAYSEATVNNNNLNATLEQLKMKTKNLLESIQANTD 150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      75  IFSKAKNLGLYDYKKTQNTLTAKAYDAATVLDNSSSTQNTNEAKTRLETAIRPAAS 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      151 KTFEDNEHPNVVEAYKALKTTLEQRAFNLGLSLSTAYNQIRNNLVDPYNNKASLLTKTLD 210
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      135 KQITDEQAAELVYKKEIKTLLSNETATLAPYDAQVAGIKMHLGLYDAGKAIATTKTLE 194

```

```

Oy      211 PLNGGLDLSNETTVNENINNTL--STINEOKN-----ADALSNPFRKVIQNNQESFV 264
           | : : : : : |
Dh      195 PVEGDP-LTASAMMANNTKIVEALKDEVLINPOKKNATKLADLSLSIVKIKITVEE- 249
           | : : : : : |
Oy      265 GFETNANVOPSNYSFEVASADVPV-----NYKARRTW-NGDEPSSRIANT-- 312
           | : : : : : |
Dh      250 ---AHKQAPNAYSFVGKRYMTTELLDKQVFPMDVAKORTIFINSDEP--RSISNTPAD 304
           | : : : : : |
Oy      313 ----NSTIDVEMTYSIAGTKTKNOFSPSNQSPSTGYYPFYKLVKADANNVGLQYKLN 367
           | : : : : : |
Dh      305 GQTMQAPLNSVMWISYLSAGTKAKYLTLETTYQFSTGYIYFFYKLVNNSDQKGLGLEYKLN 364
           | : : : : : |
Oy      368 NGNVQVEEPATISTA-----NNTTANPPPAVDEIKVAKIVUSGRFGQNTIELSVPRGEG 422
           | : : : : : |
Dh      365 D-----ATKPSATTFGSDQTMNGKPPVTDINAVAKVTLANIFSQNKIEFSVPA--- 413
           | : : : : : |
Oy      423 NNMKVAPMIGNITVLSNENNAADKI 446
           | : : : : : |
Dh      414 --EKVSPMIGNMYLSSPPNNMKI 435

```

RESULT 6  
 548755  
 major surface protein (clone pmg1.5) precursor - Mycoplasma gallisepticum (fragment)  
 C/Species: Mycoplasma gallisepticum  
 C/Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
 C/Accession: S48755  
 R/Marshall, P.F.; Grew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocke, T.D.; Browning, G.F.;  
 FEBS Lett. 352, 347-352, 1994  
 A/Title: The organisation of the multigene family which encodes the major cell surface p  
 A/Reference number: S48751; MUID:95010739; PMID:7925999  
 A/Accession: S48755  
 A/status: preliminary; nucleic acid sequence not shown  
 A/Molecule type: DNA  
 A/Residue: 1-386 #MAR>  
 A/Cross-reference: EMBL:128424; NID:9535687; PIDN:AA62419.1; PID:9535692  
 A/Note: the sequence of residues 385-386 and the corresponding nucleotide sequence are nc  
 C/genetics:  
 A/genetic code: SGC3  
 A/start codon: GTG

Query Match	24.9%	Score 579;	DB 2;	Length 386;
Best Local Similarity	38.4%	Pred. No. 1.3e-24;		
Matches 144;	Conservative 53;	Mismatches 122;	Indels 56;	Gaps 11;
QY	53	VGSTVRLRFGCSIT-----KCDANPN-----	75	
DB	15	IGSFVMLAASCTSTPTPTNHEPKPAPKPKPPGGMGMGNNGNTNPGNGMD	74	
QY	76	NGQTGLEAPMELTDLINAKMTLASLDYAKTEASVSEAEVTNNINATLEOLKM	135	
DB	75	NSAQQLAAAKKELSDLATONSNLSTYADAKINDULTAATTTAFASQNGAATLEQYKN	134	
QY	136	AKNNLESAINQANTDKTTPNEHEPNLVEAYKALKTTLEQRYNLEGLSTF--YNOIRNKL	194	
DB	135	AASTLOTPIATYAVNEKVPENNSSELVTAYTNLKTTEGENTTLAAFNDSANYGIGIKTHL	194	
QY	195	VDLYNKASSLTITKLPDLN--GGTLLDSNEITTVNRINNTL--STINQKNTADALNSF	251	
DB	195	LSLTNQAKITTTSTL--LNDAGOSPKNQNVKINKEITDAINPTILNQOKANADMLANSF	252	
QY	252	IKKYIONNEGSFVG---FTTNANQPSNYSFVARSADVTP-----VNYKYARTVWNGD	302	
DB	253	TKQVL--NDAQLTSSSSEFSMOTQPOQNGYSFVGYSVDVTGSSNNARPMWNFQORKYWDTN	311	
QY	303	EPSSRIILANTNSITDVSWITSLAGTNTKYQSFNSYGS--TGYLTFPYKLYVAAALNNGV	361	
DB	312	RAPLAQTEQSNKLTIDVSWITSLSGWGAKYTVTFDYGGASNNAYLTFPYKLVQVOTND--NYG	369	
QY	362	LOYKLNNNGVQVEF	376	
DB	370	LOYVLNNTTKLVNF	384	



Db 465 NMVLTSSDRDVNK 477

## RESULT 2

548753

major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48753

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48753; MUID:95010739; PMID:7925999

A:Accession: S48753

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 &lt;MAR&gt;

A:Cross-references: EMBL:L28424; NID:9535687; PIDN:AAA62417.1; PID:9535690

C:Genetics:

A:Genetic code: SGC3

A:start codon: GTG

Query Match 32.3%; Score 750; DB 2; Length 702;  
Best Local Similarity 41.4%; Pred. No. 1.4e-33;  
Matches 175; Conservative 62; Mismatches 126; Indels 58; Gaps 11;

QY 73 NNNNGCT---OLEAARHELTDLINAKAMTLASLDYAKIEASLSAYSEAEVNNNNNA 128

Db 69 NNNPNTPEEQOAAARKTLTDLGTEENTNVALVADYAKIOSTLSTAYMTAKTASENTSA 128

QY 129 TLEOLKMAKTNLESAINQANTDKTFEDNEHPLVYAYALKTTLEQRATNLEGSSTAYN 188

Db 129 TLENRASATTLQAAIDKANDKRVFDSYNQPLVAAINNKTTLSKTTSLGSESKYK 188

QY 189 QIRNNLVLDYNKASSLITKTLDPNG--GTLDSNEITVNNINNTLS--TINEOKTNA 244

Db 189 GIKHNLSTLFTGSAITAKTLDPGSEPTLEKVN---ANNIGIMASISPELSKKMKGNA 245

QY 245 DALNSFIKVIQNNQOSVGTFTNANQPSNYSVASADVTP-----VNYKARRTW 299

Db 246 DKF-NEFENPLSKERLSTSDTANQEPAMSPAAVSDLTNSOMLPNNFQORRW 304

QY 300 NCD--EPSRLIANTN--STDVSMIYSLAGTNTKQFSPSNYGPSTGYLYPYKLVKAD 356

Db 305 TSENOQPKTALVSESPVATDVSMIYSLAGSGTKTLTFEYIGPDMAFLYLYKLVKAD 364

QY 357 ANNVLQYKLNNGNVQVEF-----ATST-----SAN 383

Db 365 SSSVALQVSLNKTSSKLIFKPAETVSTNTDQSENEVATSTTEARSSYKVLVADEAATS 424

QY 384 NTTANPTPAVDIETAKIYLSGLRFGQNTIELSVPTGEGNNAKVAPMIGNIYLSNENNA 443

Db 425 NNEEMHTPTVSDINIAKVTLSGLTFEGENTIEFSVEG-----KVAPMIGNMVLTSNSSQ 479

QY 444 DK 446

Db 480 VKI 482

RESULT 3

548751

major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum  
N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 &lt;MAR&gt;

A:Cross-references: GB:U90714; EMBL:L28423; NID:91905869; PIDN:AA50152.1; PID:91905870

R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whitehead, K.G.

Infect. Immun. 60, 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepticum

A:Reference number: A44793; MUID:92363591; PMID:1379991

A:Contents: 56

A:Accession: A44793

A:Status: preliminary

A:Molecule type: protein

A:Residues: 26-42 &lt;MA2&gt;

A>Note: sequence extracted from NCBI backbone (NCBI:111017)

C:Genetics:

A:Genetic code: SGC3

A:start codon: GTG

Query Match 32.1%; Score 745; DB 2; Length 650;  
Best Local Similarity 40.5%; Pred. No. 2.4e-33;  
Matches 177; Conservative 65; Mismatches 139; Indels 56; Gaps 12;

QY 53 VGSVTIRLEFGCHSITKDA-----NNNGQ-----TOLEAARHELT 89

Db 15 IGSFVMLAASCTTPSPAPNPSPNGNGNINPGGCGMMAAAGELAAARWGLT 74

QY 90 DLINAKAMTLASLDYAKIEASLSAYSEAEVNNNNLNTLEOLKMAKTNLESAINQANT 149

Db 75 TVFDSKAKNLTGLYVDYKKTONTLTAKYDAKATVDSSSTTONLEAKTLETRIRIAT 134

QY 150 DKTFEDNEHPLVYAYALKTTLEQRATNLEGSSTAYNQRNNLVLDYNKASSLITKTL 209

Db 135 SKQFFDSQHAELVYKVELKTTLSNEFATLAPADAQYAGIKHNLSTLGLYDAGKAITKTL 194

QY 210 DPLNGTGLDSNEITTVNRINNTL--STINEOKTADALNSFIKVIQNNQOSFVGT 267

Db 195 EPEVGD--LTAGAVTANNTKIVEAIKDEVLPKKNATKLDSPVKVLKEKITVEEA 253

QY 268 TNANVQPSNYSPVAFSADVTPV-----NRYKARRTW--NGDEPSRLIANT----- 312

Db 254 HN-KAOPANYSFVGYSDITGTANGOTSIPNNYAOGITTTNGDEP--RVSNTPVVGQR 310

QY 313 --NSITDVSMIYSLAGTNTKQFSPSNYGPSTGYLYPYKLVKADANNVGLQYKLNNGN 370

Db 311 MAQPLSNVSMIYSLAGCAKYLEFTYGGSTGLVLYPYKLVNTSDQMKGLLEYKLANDA- 369

QY 371 VQVVEPFTSANTTN--PTPAVDIETAKIYLSGLRFGQNTIELSVPTGEGNNAKVA 429

Db 370 ---TEPSAITFGNEQTNGKTPTVNDINAKVTLANLIFGSKNIEFSVPA-----EKVSP 421

QY 430 MIGNIYLSNENNAK 446

Db 422 MIGNIYLSSEPNMVKI 438

RESULT 4

548752

major surface protein (clone pmGAL.2) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48752

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S48752

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-649 &lt;MAR&gt;

A:Cross-references: EMBL:L28424; NID:9535687; PIDN:AAA62416.1; PID:9535689

C:Genetics:

A:Genetic code: SGC3

A:start codon: GTG

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:18:32 ; Search time 42 Seconds  
(without alignments)  
1044.117 Million cell updates/sec

Title: US-09-901-572A-3

Perfect score: 2324

Sequence: 1 MHYFRNCIFPLIVLYGTN.....SSNNENADKIPGRPCTFL 456

Scoring table: BLOSUM62

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	35.6	702	2 S48754	major surface prot
2	750	32.3	702	2 S48753	major surface prot
3	745	32.1	650	2 S48751	major surface prot
4	741.5	31.9	649	2 S48752	major surface prot
5	711.5	30.6	647	2 A49218	hemagglutinin homo
6	579	23.9	386	2 S48755	major surface prot
7	314	13.5	865	1 VGBRRB	glycoprotein B pre
8	189.5	8.2	320	2 S51560	major surface prot
9	176.5	7.6	933	2 S41539	fibrogen-binding
10	170.5	7.3	4688	2 F82885	hypothetical prote
11	170	7.3	1302	1 JC6009	surface-located me
12	170	7.3	6713	2 B89921	hypothetical prote
13	166.5	7.2	3890	2 C89921	hypothetical prote
14	166	7.1	661	2 AG2422	hypothetical prote
15	164.5	7.1	2481	2 D90011	FmbB protein (impo
16	162.5	7.0	1237	2 D71850	probable outer mem
17	157.5	6.8	1072	2 A86827	cell surface anti
18	156	6.7	820	2 T17519	hypothetical prote
19	155	6.7	807	2 B71605	hypothetical prote
20	155	6.7	989	2 D89852	lmp1 protein - Myc
21	152.5	6.6	1365	2 T30822	toxoid-like outer m
22	152.5	6.6	2399	2 H71879	probable invasin Z
23	151	6.5	2660	2 E85822	subtilisin-like pr
24	150.5	6.5	682	2 S44131	kinesin-related pr
25	148.5	6.4	1073	2 S14032	surface membrane p
26	148	6.4	624	2 PC6003	ser-lys rich hypot
27	147.5	6.3	751	2 T40462	kinesin-like prote
28	147.5	6.3	1085	2 T38378	glycoprotein Vp260
29	147	6.3	1335	2 T17508	

30	146.5	6.3	1051	2 T18351	lmp1 protein - Myc
31	144.5	6.2	3194	2 D71917	toxoid-like outer m
32	144	6.2	1107	2 AC0976	probable autotrans
33	143.5	6.2	589	2 B97806	hypothetical prote
34	143	6.2	1487	2 AG2560	hypothetical prote
35	142	6.1	135	2 B49218	hemagglutinin homo
36	142	6.1	719	2 S55119	hypothetical prote
37	142	6.1	868	2 G71691	hypothetical prote
38	141.5	6.1	796	2 T21460	hypothetical prote
39	141.5	6.1	1524	2 S68553	surface layer prot
40	141	6.1	1314	1 TMBIR6	transcription regu
41	139.5	6.0	926	2 AE1130	conserved hypotet
42	139.5	6.0	1116	2 D97001	probable membrane
43	139.5	6.0	1238	2 A64596	hypothetical prote
44	137.5	5.9	568	2 E97066	membrane associate
45	137.5	5.9	1018	2 A32192	fibronectin-bindin

ALIGNMENTS

RESULT 1

S48754 major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_rev150103-Nov-1995 #ext\_change 07-Dec-1999  
C:Accession: S48754  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; V  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organization of the multigene family which encodes the major cell surface pr  
A:Reference number: S48751; PMID:95010739; PMID:7925999  
A:Accession: S48754  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-702 <MAR>  
A:Cross-references: EMBL:128424  
C:Genetics:  
A:Genetic code: GCG3  
A:Start codon: GTG

Query Match 35.6%; Score 828; DB 2; Length 702;

Best local similarity 40.6%; Pred. No. 7.9e-38;

Matches 176; Conservative 86; Mismatches 125; Indels 46; Gaps 9;

QY	50	PPVGSIVIRLIERGCSITKQANPNNG-----OTQLEAKMELTDLNAKMTLASIQ	103
DB	54	PNPGGMMGMMGG-----NTNPGGCGTDNAQQLAARKKLSDLATONSLSIVA	106
QY	104	DYAKIEASLSAYSEAEETVNNNNLNALEOLKMAKTLESAINQANTDXTPEDEHPVLVE	163
DB	107	DYANIQNTLTAAVYTAKSTSDNTSALIEOVKASITSLQTHIDTAASKISFDEKNPELIK	166
QY	164	AYKAKTTLLEQATNLEGLSTAYNOIRNNLVDIRYKASLLITKTDPLNGCTLLDSNEI	223
DB	167	AYNKLKTLKMKRNSLSGLTDSNFATIKNTLTALYQGXIVTKTDLPLM-GTAINISAV	225
QY	224	TVVKNINNTLSTNEQKTNADALNSFIKVIQNNQSGVGFETNANVOPSNYFAFS	283
DB	226	SOANTNISNAVSKELETKNTATVATLATSFKVEVLKNTLGTDT-TNNQOQPGNYSFGYS	284
QY	284	ADVTP-----VYKVARRTVMNGD-----EPSSRIIANTNSITDVSWISLACTNTRYQ	332
DB	285	VDTVTGSGNAPNWSFAQRKWTSTNTDLSQPPAEGENQGSADVSWIYVLTGMGAKYS	344
QY	333	FSFSNYGPGSTGYLYPYKLVKAADANNVGLQKLANNGVQOVERATS-----TSAN	383
DB	345	LTENVYGPSTGFLYFPYLVNNSSDSKVALEYKKNESAVKTIIDSPQTSFVADATREN	404
QY	384	N-----TTNPPFAVDEIVAKIVLSGRFGONTIELSVP-TGEQNMNVAPMIG	432
DB	405	NRSTAAPQSGSTEINPALTLDIDIKAKVTLNSLAFSGSTIEFSPTTAKBQTSVAPMIG	464
QY	433	NIYLSNENNADK 445	

TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5835

Query Match 6.7%; Score 156; DB 9; Length 2434;  
Best Local Similarity 22.9%; Pred. No. 0.0033;  
Matches 98; Conservative 68; Mismatches 164; Indels 98; Gaps 20;

```

QY 79 TQLEAARHEDLDLNAKAMTLASLDYAKIEASLSAYSEAEFTVNNNLNATLEQLKMAKT 138
DB 1003 SELNTPAMSNLQNGINDEAATKAA-QKYTDADREKOTAYNDAVTAATAKTLDDKTAGSNDKKA 1061
QY 139 NLESAINQANTDKTTFD-----NEHPN-----LVEAYKALKTTLEQRATNLEGL 182
DB 1062 AVEQALQVRNTAKTLNGDERLNEAKTRAKQOVATMSHLTDQKXANLTSQIESGTTVAGV 1121
QY 183 S-----STAYNQIRNNL-----VDLYN-----KASLITKTLD 210
DB 1122 QGIGANAGTLDQAMNQLROSIAKSDATKSEEDYQDANADLQAVYNDVATNAEGTISATNN 1181
QY 211 PLNGGTLDSNEITTVNRINNTLSTIN--EQKTADALSNSEFIKKVIQ--NNEQSFVGTG 267
DB 1182 P-----EMNPDTINQKASQVNSAKSALNGDEKLAQAKOTAKSDIGRLTDLNNAQR--TA 1233
QY 268 TNANV-QPSNYSFV-----AFSADVTPVNYKYA-----RRTV--WNGDEPSSRIILAN 311
DB 1234 ANAEVDQAPNLAAYTRAARKAKTSINTAMGNLKHALLAEKDNTRKRSVNYTDDAQPKQ--AY 1291
QY 312 TNSITDVSWIYSLAGTN--TRYQFSFSNYGPGSTGYLPYKLVKADANNVGLQYKLN 368
DB 1292 DTAAVTQAEKITVANGSNANETQVQAALQNLQAKNDLNGDNKVAQAKESAKRALASYSNL 1351
QY 369 GNVQ-----QVEFATS-----TSANNTANPTPAVDEIKVAKIVUSGLRFGONTIELSVP 418
DB 1352 NNAQSTATTSQIDNATVAGVTAQNTANEINTAMGQLQ-----NGIN-DQNTVKQOVN 1404
QY 419 TSEGNNNK 426
DB 1405 FTDADQK 1412

```

Search completed: August 14, 2003, 10:25:26  
Job time : 29 secs

```

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13083
; LENGTH: 1048
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
; US-09-815-242-13083

Query Match
Best Local Similarity 23.8%; Score 159; DB 9; Length 1048;
Best Local Similarity 23.8%; Pred. No. 0.00055;
Matches 88; Conservative 52; Mismatches 147; Indels 82; Gaps 19;

QY 79 TQLEAARMELTDLINAKMTIASLODYAKIEASISGAYSEA-ETVNNINATLEBQLKAK 137
DB 85 TDLNTMGMLQGAINDGQTLNS-QNYQATPSKATAYNAQAADILINKNGQNK-TK 142
QY 138 TNLSEAINQANTDKTTFDEHEPMLVEAYYALKTLEBQATNLEGLSTAYNOIRNNVLDL 197
DB 143 DQYTEAMNQVNSAKNNLDG---TRLDOAKQTAQOANNMTHLTQAKTMLTQO--- 193
QY 198 YNKASSLITKTLPLNGGTLIDSNELTYNNRN-INNTLSTINBOKNNADA--LSNSFI-- 252
DB 194 -----INSGLTVAGVQVQSANLTDQAMTTLRQSLANKQATASSEDYDA 239
QY 253 ---KKVIONNEQSFVGTNANVQPS-NYSFVAFSADVTVPVYKARRTVMNGDE----- 303
DB 240 NNDQOTRYNNAVAALERTIYANSPENNPSTIQKKE-QVN---SKTALNGDEMLTA 294
QY 304 -PSSRIANT-NSITD-----VSWIYS---LAGNTKIQPSNSNGPSTGYIIFYKLV 352
DB 295 KQNAKTYLNTLTSITDAQNNLISQILTSATRVSGVDITVKQNA----- 336
QY 353 KAADANNVGYQYKLNNGVQOVBEPATSTSANNTTANPTPAVDE-IVAKIVL---SGLRF 408
DB 337 QHLDQAMASLQNGINN---ESQVKSSEKTRDADITKQGEYDAITRAKAILINKSGTGPMT 392
QY 409 GQNTIELSV 417
DB 393 AQNNAVEAAL 401

RESULT 14
US-09-820-843A-108
; Sequence 108, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
; FILE REFERENCE: O63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 807
; TYPE: PRF

```

```

; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc feature
; OTHER INFORMATION: gi|3845292
; US-09-820-843A-108

Query Match
Best Local Similarity 23.8%; Score 156; DB 11; Length 807;
Best Local Similarity 23.8%; Pred. No. 0.00066;
Matches 73; Conservative 48; Mismatches 124; Indels 62; Gaps 14;

QY 122 VNNINATLEBQKMAKT-----NLESAINQA--NTDK-----TFEDNEHPNLVEAYKALT 170
DB 468 INNINISYT--QAKNNNNINFINHESPINQCHNNTEFKVNNNDPNEPNTNKKKKKEKK 525
QY 171 TLEBQATNLEGLSTAYNOIRNNVLDLYNKASSLITKTLPLNGGTLIDSNELTYNNRN 230
DB 526 NIHFNNNNNNNNKKKLYKDINQ---DHNSLIINTNQFDDH-----NNVAKTEQWL 574
QY 231 ---NNTLSTINBOKNNADALNSFIKVIQNNNEQSFVGTNANVQPSNYSFVAFSADVT 287
DB 575 QKHNNPQSVQKSNKNNKNSHLKKQJNIN-----TNNNDKNNSHISKVIYD 626
QY 288 PVNYK-----YARITVWNGDEPSS---RLANTNSITDVSWIYSLAGNTYKQ-FSPSN 337
DB 627 DNLLKSHADNSMEIYTKGKKKNTKKKKKINNINSVNNNNINNSNNIISMN 686
QY 338 YGPSTGYLFFPYLVAAANNVGLQYK-----LNNGNVQOVBEPATSTSAN 383
DB 687 VNNMNNPMPFPVNVIO-KQDSNIALLYNNKPNIDFNNFQLNHINNMIOQNNIMTNVMLN 745
QY 384 N--TTAN 388
DB 746 NNLTTSN 752

RESULT 15
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; Patent No. US20020061699A1
; GENERAL INFORMATION:
; APPLICANT: Haesideck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5835
; LENGTH: 2434

```

PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 12996  
LENGTH: 6281  
TYPE: PRF  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12996

Query Match 7.3%; Score 169.5; DB 9; Length 6281;  
Best Local Similarity 21.9%; Pred. No. 0.0011; Indels 97; Gaps 16;

Matches 96; Conservative 67; Mismatches 179; Indels 97; Gaps 16;

QY 75 NNGOTLEARMELTDLINAKAMTLASLDYAKIEASISAYSEAET-----VNNNLNAT 129  
DB 266 SNTSQTQMTAMANTLNGINDKNTILAS-ENTHDADSDKKTATQAVTAENITLNNSSN 324

QY 130 LEOLMAKTNIESAINQANTDTPDNEHPNIVEAYKALKTTLEQRATNLEGLSTAVNQ 189  
DB 325 LD-----KTAVERNALSQVANAKALNGNH-NLEQAKSNANTTI-----NGLQHLTTAOKDX 374

QY 190 IRNNLVLDLYNKS-SLITITLDELNG--GTLLDSSEITTVNRRNINNTLSTINEQCT----- 242  
DB 375 LKQVQQAQONAVGVDTVSSANTLNGAMGTLSNLSIQDNTATNGQNYLDATERNKTNYNN 434

QY 243 -----NADALNSFIKVIYQNNESFVGTFTNANQPSYSEFASADVT 287  
DB 435 AVDSANGVINATSNRMDNAINOIAITQVTSKMLDGHNTLOAKQT-----AINALDQ 489

QY 288 PVNYKARRTVNGDEPSSRIANTNSITDVSIMYSLA-----GNTTKYQSF 335  
DB 490 ATNLNAQKDALKAQVTSQAVANTSLIQDTANEINTAMGQLOHGIIDENATKOTOKYRD 549

QY 336 SNYSGTGLYFPYLYKRAADA-----NNGLOYKLN-----NGNQOYEF 376  
DB 550 AQOSKKTAY-----DQAVAAAKAILNKQGTGNSDKAAVDRALQOQVSTYKDALNGDAKLA 605

QY 377 ATSTSAN-----NTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNAN----- 425  
DB 606 KAAAKQNLGTNLHITNAQRTDLEGOINQATV-----DQVNTVKTNNANTLDGANNLSQGS 660

QY 426 ---KVAPMIGNIYLSNEN 441  
DB 661 INDKDATTLENNQVYLDDES 679

## RESULT 12

US-09-815-242-5885  
Sequence 5885, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 5885  
LENGTH: 1029  
TYPE: PRF  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5885

Query Match 6.8%; Score 159; DB 9; Length 1029;  
Best Local Similarity 23.8%; Pred. No. 0.00054; Indels 82; Gaps 19;

Matches 88; Conservative 52; Mismatches 147; Indels 82; Gaps 19;

QY 79 TQLEARMELTDLINAKAMTLASLDYAKIEASISAYSEA-ETVNNNLNANTLEOLMAK 137  
DB 73 TDLTAMGNLQGAINDQOTILNS-QNYQDAPTSKKTATVYNAVQAKDILKNSQONK-TK 130

QY 138 TNLESAINQANTDTPDNEHPNIVEAYKALKTTLEQRATNLEGLSTAVNQIRNNLVLD 197  
DB 131 DQVEANQVNSAQNNDG-----TRLLDQAKOTAKQOLNMTHTLITTAQNTLNQ----- 181

QY 198 YNKASLITITLDELNGTLLDSNEITTVNRN-INNTLSTINEQKTNADA--LSNSFT-- 252  
DB 182 -----INGTIVAGQVQVSANNTLDAQMNTLRQSIAKDKATKASSEDVDA 227

QY 253 ---KKVIQNNESFVGTFTNANQPS-NYSFVAFSADVTPVNYKARRTVNGDE----- 303  
DB 228 NNDQATYNNVAALAEITINANSNPNPSTIQKAE--QVN--SSKTLNGDENLTA 282

QY 304 -PSSRIANT-NSITD-----VSMIYS---LGTNKKYQSPSPNPSGYLYFPYKLV 352  
DB 283 KQNKTYLNTLTSITDQKNNLISQITSATRVSGVDVTKONA----- 324

QY 353 KAADANNVGLQYKLNQVQVEFATSTANNTTANPTPAVDE-IKVAKIVL---SGLRF 408  
DB 325 QHLDQAMASLQNGINN-----ESQVKSSEKRYRDAOTNKGQEYDVAITAAKAILNKSTGPT 380

QY 409 QNTTIELSV 417  
DB 381 AQNAVEAL 389

## RESULT 13

US-09-815-242-13083  
Sequence 13083, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A

;  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ;  
 ; PRIOR FILING DATE: 2000-03-21  
 ;

US-08-781-986A-5249

Query Match 7.6%; Score 176.5; DB 8; Length 936;  
Best Local Similarity 22.5%; Pred. No. 1.9e-05;  
Matches 111; Conservative 64; Mismatches 187; Indels 131; Gaps 22;

QY 2 HYFRNCJFLLVIL-----VGTNSS-----PSTONTREVVSVOLSEESTFYLCPP 52  
DB 18 HAIRKSGVAVSVLTGLIGFLSSKEADSENSVTOSDSMSNSKSDSSVSAAKT 77  
QY 53 VGSIVIRLEFGCMSITTKDANPNNGOTOL--EAAAMELTDLINAKAMTLASIDYAKTEA 110  
DB 78 DDIVN-----SDTKTSSNTNNGETSVAGNPAGQETTOSSTNATL----- 117  
QY 111 SLSASVSEAFYNNNLNLTLEQLKAKTNLESAIQAQANTDKTPNEHPNLVEAKAKT 170  
DB 118 EETPTGATTTTNTQAANTPATTOSSNTNAEELVNGTSETTSNDT---NTVSSVNS--- 171  
QY 171 TLEQRTNLEGLST-----AYNQIRNLLVDLYNK-----ASSLIT 206  
DB 172 --PQNSTAENVSTTQDSTTEATPSNBSAPOSTDASNKDVNOAVNTSAPMRAFLAA 229  
QY 207 KTLDELNGTLLDSNEITTVNNINNTLSTINEQ---KTN-ADALSNSFIK---KYIQ 257  
DB 230 VAADAPVAGTDI--TQLTNTVTVGIDSGTTVPYHQAQYKLVNGFVPSNAVKGDTFKITV 288  
QY 258 NNEQSFVGTFTNANQPSNYSFVAFSADVTPTNYKARTVWNGDEPSSRIANTNSITD 317  
DB 289 PEELNLNGVTSTAKVP-----IMAGQ---VLANGVITDSD 321  
QY 318 VSMIYSLAG--TWTKYQFSFSNYGSTGYLYFPYKLVKADANNVGLQYKLVNGVNGVQVEF 376  
DB 322 GNVITTFPDYNTKODVAKITLMPA--YI-----DPENV-----KKTGNV----- 359  
QY 377 ATSTSANNTANPTPAVDEIKYAKI-----VLSGLRGQNTIELSVPTGEGNNKVP 429  
DB 360 TLATIGSTTANKTVLVDEYKGYKFNLSIKGTIDIDKTNNTYQTIYVNSGDNVLAIP 419  
QY 430 MI-GNIIYLSNEN 441  
DB 420 VLTGNLKPVTDSN 432

## RESULT 8

US-10-056-052-2  
Sequence 2, Application US/10056052  
Publication No. US2003009656A1  
GENERAL INFORMATION:  
APPLICANT: PATTI, Joseph M  
APPLICANT: HUTCHINS, Jeff T  
APPLICANT: DOMANSKI, Paul  
APPLICANT: PATEL, Pratiksha  
APPLICANT: HALL, Andrea  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN  
FILE REFERENCE: P0706US04/BAS  
CURRENT APPLICATION NUMBER: US/10/056,052  
CURRENT FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/308,116  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/298,413  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/274,611  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: 60/264,072  
PRIOR FILING DATE: 2001-01-26  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patent version 3.1  
SEQ ID NO 2  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-10-056-052-2

Query Match 7.4%; Score 173; DB 15; Length 520;  
Best Local Similarity 22.7%; Pred. No. 1.5e-05;  
Matches 105; Conservative 62; Mismatches 173; Indels 122; Gaps 21;

QY 24 STQNTSREVVSVOLSEESTFYLCPPVGSIVIRLEFGCMSITTKDANPNNGOTOL-- 81  
DB 1 SENSVTOSDSMSNSKSDSSVSAAPKTDITNV-----SDTKTSSNTNNGETSVAG 52  
QY 82 EAAAMELTDLINAKAMTLASIDYAKTEBASLSASVSEAFYNNNLNLTLEQLKAKTNLE 141  
DB 53 NPAQOETTOSSTNATL-----EETPTGATTTTNTQAANTPATTOSSNTNAE 100  
QY 142 SAIQQAQNTDKTPNEHPNLVEAKAKTLEQRTNLEGLST-----AYNQIRN 192  
DB 101 ELVNG--TSNETTFD--TNTVSSVNS-----PQNSTAENVSTTQDSTTEATPSNBSAP 152  
QY 193 NLVDLYNK-----ASSLITKTLDELNGTLLDSNEITTVNNINNTLSTI 237  
DB 153 QSTDASNKDVNOAVNTSAPMRAFLAAVAADAPAGTDI--TQLTNTVTVGIDSGTTVY 211  
QY 238 NEQ---KTN-ADALSNSFIK---KYIQNNEQSFVGTFTNANQPSNYSFVAFSADVT 288  
DB 212 PHQAQYKLVNGFVPSNAVKGDTFKITVPEELNLNGVTSTAKVP----- 257  
QY 289 VNYKARTVWNGDEPSSRIANTNSITDVSMIYSLAG--TWTKYQFSFSNYGSTGYLYF 347  
DB 258 -----IMAGQ---VLANGVITDSDGNVITTFPDYNTKODVAKITLMPA--YI-- 300  
QY 348 PYKLVKADANNVGLQYKLVNGVNGVQVEFATSTSANNTANPTPAVDEIKYAKI----- 401  
DB 301 -----DPENV-----KKTGNV-----TLATIGSTTANKTVLVDEYKGYKFNLSIK 342  
QY 402 -VLSGLRGQNTIELSVPTGEGNNKVPML-GNIIYLSNEN 441  
DB 343 GTIDQIDKTNNTYQTIYVNSGDNVLAIPVLTGNLKPVTDSN 384

## RESULT 9

US-09-815-242-5639  
Sequence 5639, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyckind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5639  
LENGTH: 2086

SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 23  
 LENGTH: 357  
 TYPE: PRT  
 ORGANISM: Mycoplasma gallisepticum  
 FEATURE: Modified TTT-1 portion (downstream of Bgl1) of  
 OTHER INFORMATION: pMZ40K-S  
 US-10-131-591A-23

Query Match 76.2%; Score 1770; DB 15; Length 357;  
 Best Local Similarity 98.0%; Pred. No. 3.6e-133;  
 Matches 350; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 89 TDLINAKAMTLASIQDYAKIEASISAYSEAEVTNNNNIATLEQAKAKTLESAIQAQ 148  
 DB 1 TDLINAKAMTLASIQDYAKIEASISAYSEAEVTNNNNIATLEQAKAKTLESAIQAQ 60  
 QY 149 TDKTFENDEHPVLVAYKALKTTLEQRADNLEGLSTAYNOIRNNLVLYNKASSLITKT 208  
 DB 61 TDKTFENDEHPVLVAYKALKTTLEQRADNLEGLSTAYNOIRNNLVLYNKASSLITKT 120  
 QY 209 LDPLNGTLLDSNEITTVNRNINNTLSTINEQKTNADALSNSFIKVIQNNQSFVGTFT 268  
 DB 121 LDPLNGTLLDSNEITTVNRNINNTLSTINEQKTNADALSNSFIKVIQNNQSFVGTFT 180  
 QY 269 NANVOPSNVGFVAFSADVTPVNYKARRTWNGDEPSSRIILANTNSITDVSMIYSLAGTN 328  
 DB 181 NANVOPSNVGFVAFSADVTPVNYKARRTWNGDEPSSRIILANTNSITDVSMIYSLAGTN 240  
 QY 329 TKYOPSEFNSYSGTGLVPEPYKLVKAADANNVGLQYKLNNGVQVEPATSTANNTAN 388  
 DB 241 TKYOPSEFNSYSGTGLVPEPYKLVKAADANNVGLQYKLNNGVQVEPATSTANNTAN 300  
 QY 389 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGMNNKVPAMIGNIYISSNENNDK 445  
 DB 301 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGMNNKVPAMIGNIYISSNENNDK 357

RESULT 5  
 US-10-131-591A-5  
 Sequence 5, Application US/10131591A  
 Publication No. US20030059799A1  
 GENERAL INFORMATION:  
 APPLICANT: Nippon Zeon Co., Ltd.  
 TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
 FILE REFERENCE: J209  
 CURRENT APPLICATION NUMBER: US/10/131,591A  
 CURRENT FILING DATE: 2002-08-15  
 NUMBER OF SEQ ID NOS: 79  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 62  
 TYPE: PRT  
 ORGANISM: Marek's disease gammaherpesvirus  
 FEATURE:  
 OTHER INFORMATION: MDVgB signal  
 US-10-131-591A-5

Query Match 13.8%; Score 320; DB 15; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-18;  
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILYIGTSSPSTONVTSREYVSSVQSEESTFYLCPPVGSIVIRL 60  
 DB 1 MHYFRNCIFFLVILYIGTSSPSTONVTSREYVSSVQSEESTFYLCPPVGSIVIRL 60  
 QY 61 EF 62  
 DB 61 EF 62

RESULT 6

US-10-131-591A-6  
 Sequence 6, Application US/10131591A  
 Publication No. US20030059799A1  
 GENERAL INFORMATION:  
 APPLICANT: Nippon Zeon Co., Ltd.  
 TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
 FILE REFERENCE: J209  
 CURRENT APPLICATION NUMBER: US/10/131,591A  
 CURRENT FILING DATE: 2002-08-15  
 NUMBER OF SEQ ID NOS: 79  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 62  
 TYPE: PRT  
 ORGANISM: Marek's disease gammaherpesvirus  
 FEATURE:  
 OTHER INFORMATION: Modified VgB signal  
 US-10-131-591A-6

Query Match 13.3%; Score 308; DB 15; Length 62;  
 Best Local Similarity 96.8%; Pred. No. 1.1e-17;  
 Matches 60; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILYIGTSSPSTONVTSREYVSSVQSEESTFYLCPPVGSIVIRL 60  
 DB 1 MHYFRNCIFFLVILYIGTSSPSTONVTSREYVSSVQSEESTFYLCPPVGSIVIRL 60  
 QY 61 EF 62  
 DB 61 EF 62

RESULT 7  
 US-08-781-986A-5249  
 Sequence 5249, Application US/08781986A  
 Publication No. US2003005436A1  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30,446  
 REFERENCE/DOCKET NUMBER: PB248PP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 5249:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 936 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein



```

Db      121 TNNNNLNTLEQLQAKNTLSAIOANTDKTTPNEHNPVLEAKKLTLEQRATYLE 180
Qy      121 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      181 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Qy      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 300
Db      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 300
Qy      301 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      301 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Qy      361 GLOQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Db      361 GLOQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Qy      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456
Db      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456

```

# RESULT 2

```

US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A10RU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

```

Query Match 86.4%; Score 2008; DB 9; Length 1086;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-151;  
 Matches 384; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      61 EFGCSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLODYAKIEASLSAYSSEAE 120
Db      691 EFGCSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLODYAKIEASLSAYSSEAE 750
Qy      121 TNNNNLNTLEQLQAKNTLSAIOANTDKTTPNEHNPVLEAKKLTLEQRATYLE 180
Db      121 TNNNNLNTLEQLQAKNTLSAIOANTDKTTPNEHNPVLEAKKLTLEQRATYLE 180
Qy      181 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      181 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Qy      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 300
Db      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 300
Qy      301 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      301 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Qy      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456
Db      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456

```

```

Qy      361 GLOQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Db      991 GLOQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 1050
Qy      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456
Db      1051 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 1086

```

## RESULT 3

```

US-10-131-591A-12
; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
US-10-131-591A-12

```

Query Match 83.3%; Score 1936.5; DB 15; Length 384;  
 Best Local Similarity 99.7%; Pred. No. 2.1e-146;  
 Matches 384; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

Qy      61 EFGCSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLODYAKIEASLSAYSSEAE 120
Db      1 EF-CMSITKDKANPNNGQTOLEAARMELTDLINAKAMTSLASLODYAKIEASLSAYSSEAE 59
Qy      121 TNNNNLNTLEQLQAKNTLSAIOANTDKTTPNEHNPVLEAKKLTLEQRATYLE 180
Db      121 TNNNNLNTLEQLQAKNTLSAIOANTDKTTPNEHNPVLEAKKLTLEQRATYLE 119
Qy      181 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      181 GLSSTAVYQIRNNLVLDYLNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 179
Qy      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 300
Db      241 KTNADALNSFIKKYIIONNEGSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWN 239
Qy      301 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      240 GDEPSSRIANTNSITDVSWMYISLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 299
Qy      421 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 456
Db      360 EGNMNVKAPMIGNIYILSSNENNADKIPGYRRPGTFL 384

```

## RESULT 4

```

US-10-131-591A-23
; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:21:13 ; Search time 28 Seconds  
(without alignments)  
2133.482 Million cell updates/sec

Title: US-09-901-572a-3  
Perfect score: 2324  
Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNNENADKIPGYRRPCTFL 456

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 492763

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2317	99.7	456	US-09-147-052-2	Sequence 2, Appl1
2	2008	86.4	1086	US-09-147-052-4	Sequence 4, Appl1
3	1936.5	83.3	384	US-10-131-591A-12	Sequence 12, Appl1
4	1770	76.2	357	US-10-131-591A-23	Sequence 23, Appl1
5	320	13.8	62	US-10-131-591A-5	Sequence 5, Appl1
6	308	13.3	62	US-10-131-591A-6	Sequence 6, Appl1
7	176.5	7.6	936	US-08-781-986A-5249	Sequence 5249, Ap
8	173	7.4	520	US-10-056-052-2	Sequence 2, Appl1
9	170	7.3	2086	US-09-815-242-5639	Sequence 5639, Ap
10	169.5	7.3	5795	US-09-815-242-12610	Sequence 12610, A
11	169.5	7.3	6281	US-09-815-242-12996	Sequence 12996, A
12	159	6.8	1029	US-09-815-242-5885	Sequence 5885, Ap
13	159	6.8	1048	US-09-815-242-13083	Sequence 13083, A
14	156	6.7	807	US-09-820-843A-108	Sequence 108, Ap
15	156	6.7	2434	US-09-815-242-5835	Sequence 5835, Ap

16	154.5	6.6	2478	US-09-815-242-5816	Sequence 5816, Ap
17	154.5	6.6	2478	US-09-815-242-12967	Sequence 12967, A
18	153.5	6.6	3158	US-09-815-242-12611	Sequence 12611, A
19	147	6.3	1021	US-09-815-242-5471	Sequence 5471, Ap
20	147	6.3	1021	US-09-815-242-12544	Sequence 12544, A
21	146.5	6.3	837	US-09-815-242-5883	Sequence 5883, Ap
22	146.5	6.3	875	US-09-815-242-13080	Sequence 13080, A
23	143	6.2	1215	US-09-815-242-5908	Sequence 5908, Ap
24	143	6.2	1269	US-09-815-242-13113	Sequence 13113, A
25	143	6.2	2437	US-09-815-242-5834	Sequence 5834, Ap
26	140.5	6.0	596	US-09-946-374-243	Sequence 243, Ap
27	140.5	6.0	596	US-10-015-387A-243	Sequence 243, Ap
28	140.5	6.0	596	US-10-063-735-100	Sequence 243, Ap
29	140.5	6.0	596	US-10-006-130A-243	Sequence 243, Ap
30	140.5	6.0	596	US-10-199-672-310	Sequence 310, Ap
31	140.5	6.0	596	US-10-006-867-100	Sequence 310, Ap
32	140.5	6.0	596	US-10-052-586-310	Sequence 310, Ap
33	140.5	6.0	596	US-10-063-5847-100	Sequence 310, Ap
34	140.5	6.0	596	US-10-174-590-310	Sequence 310, Ap
35	140.5	6.0	596	US-10-176-758-310	Sequence 310, Ap
36	140.5	6.0	596	US-10-175-737-310	Sequence 310, Ap
37	140.5	6.0	596	US-10-063-616-100	Sequence 310, Ap
38	140.5	6.0	596	US-10-173-706-310	Sequence 310, Ap
39	140.5	6.0	596	US-10-175-728-310	Sequence 310, Ap
40	140.5	6.0	596	US-10-175-752-310	Sequence 310, Ap
41	140.5	6.0	596	US-10-176-482-310	Sequence 310, Ap
42	140.5	6.0	596	US-10-176-757-310	Sequence 310, Ap
43	140.5	6.0	596	US-10-176-913-310	Sequence 310, Ap
44	140.5	6.0	596	US-10-180-552-310	Sequence 310, Ap
45	140.5	6.0	596	US-10-180-557-310	Sequence 310, Ap

## ALIGNMENTS

RESULT 1  
US-09-147-052-2  
Sequence 2, Application US/09147052  
Patent No. US20010014335A1  
GENERAL INFORMATION:  
APPLICANT: SAITOH, Shuji  
APPLICANT: TSUZAKI, Yoshihara  
APPLICANT: YANAGIDA, No. US20010014335A1  
TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,  
FILE REFERENCE: 981167  
CURRENT APPLICATION NUMBER: US/09/147,052  
CURRENT FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: JP 08-103548  
PRIOR FILING DATE: 1996-03-29  
PRIOR APPLICATION NUMBER: PCT/JP97/01084  
PRIOR FILING DATE: 1997-03-28  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 456  
TYPE: PRT  
ORGANISM: hybrid  
US-09-147-052-2  
Query Match  
Best Local Similarity 99.7%; Score 2317, DB 9, Length 456;  
Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLIVLYGTNSPPONTSEVSVOLSEESTFYLCPPVGSFVIRL 60  
1 MHYFRNCIFFLIVLYGTNSPSTONTSEVSVOLSEESTFYLCPPVGSFVIRL 60  
DB 1 MHYFRNCIFFLIVLYGTNSPSTONTSEVSVOLSEESTFYLCPPVGSFVIRL 60  
QY 61 EGGCSITKQANPNNGOTLEAARMELTDLINAAMTLASLODYAKIEASSAYSRAE 120  
61 EGGCSITKQANPNNGOTLEAARMELTDLINAAMTLASLODYAKIEASSAYSRAE 120  
DB 121 TVNNINMTLEOLKRAKTNLESAINQANTDKTPNEHPNLEAVKYLKTLLEQRATYLE 180  
QY 121 TVNNINMTLEOLKRAKTNLESAINQANTDKTPNEHPNLEAVKYLKTLLEQRATYLE 180